



UNIVERSIDAD PERUANA  
**CAYETANO HEREDIA**

FACULTAD DE CIENCIAS E INGENIERÍA

Alteración del número de copias de genes reguladores de inflamación como factor predictor de la sobrevida global en pacientes con Hepatocarcinoma e infección viral por VHB/VHC: un estudio bioinformático

Tesis para optar el Título Profesional de Licenciado en Biología

Autor:

Mary Luz Gutierrez Quilca

Asesora:

PhD. MSc Claudia Ines Gloria Machicado Rivero

Lima – Perú

2025

**Jurado calificador**

**Presidente:** Dr. Jorge Enrique Rodriguez Bailon

**Vocal:** Dr. Stephane Bertani

**Secretario:** Dr. Wilfredo Evangelista Falcon



Similitud 10% Marcas de alerta



UNIVERSIDAD PERUANA  
**CAYETANO HEREDIA**

FACULTAD DE CIENCIAS E INGENIERÍA

Alteración del número de copias de genes reguladores de inflamación como factor predictor de la supervivencia global en pacientes con Hepatocarcinoma e infección viral por VHB/VHC: un estudio bioinformático

Tesis para optar el Título Profesional de Licenciado en Biología

Autor:

Mary Luz Gutierrez Quila

Asesora:

Ph.D. MSc. Claudia Ines Gloria Machicado Rivero

Lima - Perú

2025

Página 1 de 65 12671 palabras

107%



Informe estándar  
Informe en inglés no disponible Más información

10% Similitud estándar

Filtros

2 Exclusiones

Fuentes

Mostrar las fuentes solapadas

1 Internet  
hdl.handle.net <1%  
11 bloques de texto 89 palabra que coinciden

2 Internet  
docplayer.es <1%  
3 bloques de texto 48 palabra que coinciden

3 Internet  
assets.researchsquare.com <1%  
1 bloques de bloques 43 palabra que coinciden

4 Internet  
eprints.ucm.es <1%  
4 bloques de texto 34 palabra que coinciden

## **Agradecimientos**

Finalizar esta tesis representa no solo el cierre de una etapa académica, sino también el reflejo del esfuerzo, la perseverancia y el apoyo incondicional de personas que han sido fundamentales en este camino.

En primer lugar, quiero expresar mi más sincero agradecimiento a mi asesora Claudia Machicado, por su guía, paciencia y valiosos consejos. Su conocimiento y dedicación fueron esenciales para el desarrollo de esta investigación, y su confianza en mis capacidades me motivó a dar lo mejor de mí en cada paso del proceso.

A Dios y mi familia, quienes han sido mi pilar inquebrantable. Gracias por su amor incondicional, por creer en mí incluso en los momentos en que yo misma dudé, y por brindarme el apoyo emocional y material necesario para alcanzar esta meta. Sin ustedes, este logro no sería posible.

A David por su comprensión, paciencia y palabras de aliento en los momentos más desafiantes. Su presencia ha sido una fuente constante de motivación y alegría en este proceso.

## Índice

Resumen	1
Abstract	2
I. Introducción	3
1.1. Epidemiología del Hepatocarcinoma	3
1.2. Diagnóstico y pronóstico del Hepatocarcinoma	3
1.3. Factores de Riesgo	4
1.4. Alteraciones genéticas en el HCC	5
1.5. Inflamación como mecanismo mediador del HCC	7
1.6. Aproximaciones bioinformáticas para el estudio del perfil molecular del HCC	8
1.7. Base de datos utilizado en el estudio	9
II. Planteamiento del problema	10
III. Pregunta de investigación	11
IV. Objetivos	11
4.1. Objetivo principal	12
4.2. Objetivos específicos	12
V. Materiales y métodos	12
5.1. Diseño de estudio	12
5.2. Población de estudio	12
5.3. Operacionalización de variables	14
5.4. Recursos bioinformáticos	17
5.5. Obtención de datos genómicos y metadata	19
5.5.1. Construcción de base de datos de la población	19
5.5.2. Lista de genes de interés	19
5.5.3. Obtención de la ontología de los genes de inflamación	19
5.5.4. Localización de los genes reguladores de inflamación en los diferentes cromosomas	22

5.5.5. Obtención de la ANC en la cohorte GDC TCGA-LIHC	24
5.5.6. Incorporación de la metadata y ANC	25
5.5.7. Enriquecimiento funcional de los genes reguladores de inflamación	26
5.5.8. Elaboración del plot KM y ANC	27
5.6. Análisis estadístico	31
5.7. Cuestiones éticas	35
VI. Resultados	36
6.1. Descripción clínica-demográfica de la cohorte GDC TCGA-LICH	36
6.2. Ontología y distribución cromosómica de los genes reguladores de inflamación	39
6.3. Relación entre la ANC y el estado viral	42
6.4. Asociación entre la ANC y OS en pacientes con infección viral	44
VII. Discusión	48
VIII. Conclusiones	53
IX. Limitaciones	53
X. Referencias bibliográficas	54
Anexos	60

## Resumen:

El hepatocarcinoma (HCC) es el cáncer de hígado más frecuente y agresivo, siendo la tercera causa de mortalidad por cáncer a nivel mundial. El factor etiológico principal es la infección crónica por los virus de la hepatitis B (VHB) y C (VHC), que transforma los hepatocitos y provoca una inflamación crónica. En este contexto, se ha observado que ciertos genes relacionados con la inflamación presentan sobreexpresión, afectando negativamente la sobrevida de los pacientes. Sin embargo, no está claro si la infección por VHB o VHC está relacionada con alteraciones en el número de copias de genes (ANC) que regulan la inflamación, ni cuál es el impacto de dichas alteraciones en la sobrevida global. Para investigar esto, se realizó un análisis de datos genómicos de la cohorte de HCC de The Cancer Genome Atlas (GDC TCGA - LIHC), evaluando las ANC con el recurso web CBioPortal. Los resultados revelaron alteraciones en cuatro genes específicos (*CCL24*, *NAMPT*, *SLC7A2* y *SRI*), involucrados en el progreso del HCC. Al analizar la asociación entre las ANC y la sobrevida global mediante el plot de Kaplan-Meier, se observó que los participantes con amplificaciones de genes en la vía de señalización Janus quinasas - Transductor de señales y activador de la transcripción (JAK-STAT) presentaron un peor pronóstico, sugiriendo su implicación en procesos oncogénicos del HCC. Estos hallazgos abren nuevas vías para el desarrollo de biomarcadores pronósticos y potenciales objetivos terapéuticos, que podrían mejorar el tratamiento y la sobrevida de los pacientes. Es fundamental continuar investigando estas ANC y determinar su papel exacto en el HCC.

Palabras clave: Genes reguladores de inflamación, ANC, sobrevida global, VHB/VHC

## **Abstract:**

Hepatocellular carcinoma (HCC) is the most common and aggressive liver cancer, ranking as the third leading cause of cancer-related mortality worldwide. The main etiological factor is chronic infection with hepatitis B (HBV) and C (HCV) viruses, which transform hepatocytes and induce chronic inflammation. In this context, it has been observed that certain inflammation-related genes are overexpressed, negatively affecting patient survival. However, it is unclear whether HBV or HCV infection is associated with alterations in gene copy number (CNA) regulating inflammation or the impact of such alterations on overall survival. To investigate this, a genomic data analysis of the HCC cohort from The Cancer Genome Atlas (GDC TCGA - LIHC) was performed, evaluating CNAs using the CBioPortal web resource. The results revealed alterations in four specific genes (CCL24, NAMPT, SLC7A2, and SRI) involved in the progression of HCC. When analyzing the association between CNAs and overall survival through Kaplan-Meier plots, it was observed that participants with gene amplifications in the Janus kinase - Signal Transducer and Activator of Transcription (JAK-STAT) signaling pathway had a worse prognosis, suggesting their involvement in oncogenic processes in HCC. These findings open new avenues for the development of prognostic biomarkers and potential therapeutic targets that could improve patient treatment and survival. It is crucial to continue investigating these CNAs and determine their exact role in HCC.

**Keywords:** Inflammation-regulating genes, CNA, overall survival, HBV/HCV

## **1. Introducción**

### **1.1. Epidemiología del Hepatocarcinoma**

El hepatocarcinoma (HCC) es el tipo de cáncer de hígado más común y agresivo, ocupando el tercer lugar como causa de muerte por cáncer a nivel mundial. El continente asiático presenta la mayor tasa de mortalidad e incidencia de este cáncer, con un 56% y 49% respectivamente. En el año 2022, se registraron aproximadamente 865,000 nuevos casos de cáncer de hígado en el mundo, posicionándolo como el tercer cáncer más frecuente a nivel global, siendo China el país con la tasa de mortalidad más alta por esta enfermedad[1].

En Perú, al igual que en muchas regiones del mundo, el cáncer de hígado representa un desafío cada vez mayor para la salud pública. Según el informe de GLOBOCAN hasta el año 2022, la tasa de mortalidad en Perú fue de 5 por cada 100.000 habitantes, posicionándose como la octava causa de muerte por cáncer en el país [2].

### **1.2. Diagnóstico y pronóstico del HCC**

Para la detección temprana del HCC, se utilizan pruebas de imágenes como la ecografía abdominal, la resonancia magnética y la tomografía computarizada así como también pruebas de laboratorio incluyendo el análisis de alfafetoproteína sérica (AFP). Sin embargo, las pruebas de imágenes no son suficientemente sensibles cuando las lesiones son menores a 1-2 cm, lo que a menudo resulta en la detección del HCC en etapas avanzadas [3]. Debido a esto, y a pesar de los avances científicos significativos, sigue siendo crucial identificar marcadores eficaces para el diagnóstico temprano y otros para marcadores de pronóstico y tratamiento del HCC.

En cuanto al pronóstico, este del cáncer se mide principalmente mediante la sobrevida global (OS) a 5 años y la respuesta al tratamiento anticancerígeno. Específicamente, la OS varía según el estadio en que se diagnostique y el tratamiento administrado, aunque también parece estar influenciada por el género, con las mujeres mostrando medianas de sobrevida más altas [4]. Asimismo, las tasas de supervivencia relativa a 5 años varían según la localización del tumor, siendo del 33% para la etapa localizada (cuando el cáncer se encuentra limitado al hígado), 11% para la etapa regional (cuando el cáncer se ha diseminado a tejidos o ganglios linfáticos cercanos), y 2% para la etapa distante (etapa más avanzada, en la cual el cáncer se ha propagado a otros órganos o tejidos) [5]. Por ello, la detección temprana, el manejo multidisciplinario y la prevención son fundamentales para mejorar la supervivencia en el carcinoma hepatocelular (HCC) y reducir su impacto en la salud pública.

### **1.3. Factores de riesgo**

Los factores de riesgo más significativos para el HCC incluyen la infección crónica por los virus de hepatitis B (VHB) o C (VHC), la esteatohepatitis no alcohólica, la obesidad, el tabaquismo, el alcoholismo, la diabetes tipo 2 y la ingesta de alimentos contaminados por aflatoxinas [6,7]. Las hepatitis crónicas B y C están presentes en más del 70% de los casos de carcinoma hepático primario [8]. En el Perú, la tasa de incidencia de VHB y VHC por cada 100,000 habitantes es de 1.32 y 0.25, respectivamente [9,10].

El VHB se transmite principalmente por contacto con sangre o fluidos corporales infectados, por relaciones sexuales sin protección o de madre a hijo durante el parto. Este virus muestra un tropismo tisular por las células del hígado, donde se replica. Por otro lado, el VHC suele ser asintomático y a menudo provoca una infección crónica, que puede progresar a cirrosis hepática y finalmente al HCC, convirtiéndose en un problema de salud pública. Los síntomas más frecuentes incluyen náuseas, fatiga, molestias articulares y signos de daño hepático, como ictericia y niveles elevados de enzimas hepáticas [11].

Si un individuo es infectado por el virus de hepatitis, este es reconocido por los receptores de las células hepáticas mediante su antígeno y finaliza instalándose en los hepatocitos. En el caso del VHB, una vez que el virión ingresa a la célula, se descompone y libera su ADN

viral en el núcleo de la célula hospedera, el cual se transcribe en ARN viral, que actúa como plantilla para la producción de nuevas partículas virales. Los componentes virales se ensamblan dentro del hepatocito y, posteriormente, las nuevas partículas se liberan en la sangre, donde pueden infectar más células hepáticas o ser transmitidas a otras personas [12]. En cuanto al VHC este ingresa al hepatocito mediante endocitosis. Una vez en el citoplasma, su ARN de cadena positiva se traduce en péptidos maduros, los cuales son procesados por proteasas tanto virales como del huésped. En el retículo endoplásmico, la ARN polimerasa NS5B replica el genoma viral, produciendo nuevas cadenas de ARN. Estas, posteriormente, se ensamblan en viriones maduros que son liberados por exocitosis. A diferencia del VHB, el VHC no se incorpora al genoma del huésped [13].

#### **1.4. Alteraciones genéticas en el HCC**

Las alteraciones genéticas, conocidas también como mutaciones, son cambios en la secuencia del ADN que pueden ocurrir de manera natural o ser provocados por factores externos. En el contexto del carcinoma hepatocelular, las mutaciones están asociadas con factores subyacentes como VHB, VHC, el alcoholismo, diabetes tipo 2 y la exposición alimentaria a la aflatoxina B1 (AFB1). Estas mutaciones iniciales desencadenan la inestabilidad genómica (mayor frecuencia de cambios genéticos, reordenamientos cromosómicos y aneuploidía) [14], la cual se acumula gradualmente en un conjunto limitado de genes en las primeras etapas antes del desarrollo del cáncer [15].

Tal como reporta Sohrab y colaboradores [16], es necesaria la acumulación de mutaciones genéticas en los hepatocitos para la hepatocarcinogénesis; así mismo, a medida que el HCC progresa, la acumulación de mutaciones se va acelerando en todo el genoma del individuo. Sin embargo, ciertas regiones genómicas y genes, como *TP53* y *CTNNB1*, mutan con frecuencia y afectan a 21% y 40% respectivamente de pacientes con esta neoplasia [17]. Estos genes mutados activan diversas rutas de señalización, impulsando la transformación progresiva de células sanas en cancerosas [14, 18].

En este contexto, la alteración del número de copias (ANC) es una forma de modificación genética en la que una sección de ADN puede duplicarse (comúnmente afectando a un oncogen) o eliminarse (delección, impactando a un gen supresor de tumor). Se estima que un 2-14% de los casos de carcinoma hepatocelular posee ANC. En el contexto de la infección crónica por VHB, se sabe que este virus tiene la capacidad de activar oncogenes en las células hepáticas mediante la mutagénesis por inserción viral [8]. Según el estudio de Ally et al. [19], el ADN viral del VHB se integra en el genoma del hospedero en el 84% de los casos, lo que genera inestabilidad genómica, una elevada frecuencia de mutaciones y en parte provoca ANCs [7, 17]. Por otro lado, según el estudio Zhao et. al (2021) el VHC altera vías de señalización como Wnt y p53 mediante hipermetilación y ANCs, afectando TP53, MYC y WNT. Estas modificaciones promueven la proliferación celular incontrolada, la evasión de la apoptosis y la alteración de los mecanismos de reparación del ADN. Como resultado, se favorece la acumulación de mutaciones y la transformación maligna de los hepatocitos, aumentando el riesgo de progresión hacia el cáncer hepático [20].

En particular, se observa una delección del promotor del gen de la transcriptasa inversa de la telomerasa [*TERT*] en el 60% de los casos de individuos con HCC [7]. Además, la ANC en la región 1q21 es una de las modificaciones más frecuentes en el HCC, con una incidencia del 58-86%. Esta alteración ocurre en las etapas iniciales del desarrollo tumoral, la cual ocasiona la sobreexpresión de genes, que participan en funciones como el mantenimiento de los telómeros, la regulación del ciclo celular, las vías Wnt/ $\beta$ -catenina y de la fosfatidilinositol-3 quinasa, además del control del estrés oxidativo [14]. Por ejemplo, las células hepáticas que presentan translocación de  $\beta$ -catenina al núcleo experimentan un crecimiento celular desregulado, expresan proteínas de membrana asociadas al HCC y muestran comportamiento metastásico [14]. Por tanto, la existencia de múltiples copias de genes, puede conducir a un crecimiento celular descontrolado y al surgimiento de tumores más agresivos [21].

## 1.5. Inflamación como mecanismo mediador del HCC

La inflamación se refiere a la respuesta del cuerpo frente al daño en los tejidos, el cual puede ser ocasionado por lesiones físicas, isquemia, infecciones o exposición a sustancias tóxicas. Esta respuesta desencadena cambios celulares y respuestas del sistema inmunológico que conducen a la reparación de la lesión producida. En caso fallen estos mecanismos reguladores y no logren mediar la inflamación, termina convirtiéndose en una condición crónica en el tejido afectado, donde ocurren mutaciones, así como una proliferación excesiva de células, creándose así un ambiente propicio para el desarrollo de cáncer [16, 22].

Con respecto al HCC, un 80-90% de los casos se desarrolla en un entorno de inflamación crónica causada por diversos factores como infecciones virales (VHB, VHC), consumo excesivo de alcohol y trastornos metabólicos hepáticos, como la esteatohepatitis no alcohólica o la enfermedad hepática grasa no alcohólica [23]. Esta inflamación altera el microambiente de los hepatocitos y, en el caso de infecciones virales, los virus pueden eludir la respuesta inmunitaria (tanto innata como adaptativa). En el transcurso de este proceso, se generan mutaciones que facilitan la progresión de la enfermedad, desde la inflamación hasta la formación del tumor [18].

Asimismo, la inflamación crónica se caracteriza por la infiltración de macrófagos y células mieloides, así como por una producción descontrolada de citocinas que promueven la carcinogénesis en el HCC. Durante este proceso, tanto las células inmunitarias como los hepatocitos dañados, liberan mediadores inflamatorios como citoquinas proinflamatorias (IL-6, IL-10, TNF- $\alpha$ ) y quimiocinas, que reclutan más células inflamatorias al hígado, promoviendo así la proliferación celular. Además, se genera estrés oxidativo que produce especies reactivas de oxígeno (ROS), las cuales dañan el ADN de las células hepáticas y al acumularse pueden llevar a alteraciones genómicas, con mutaciones en oncogenes y supresores de tumores [7]. En el contexto del HCC y la inflamación asociada a VHB/VHC, algunos genes como *CCL20*, *IL1B*, *CCL2*, *CCL22* y *TIMP1* muestran una sobreexpresión significativa, que fomentan el crecimiento tumoral, la invasión y la resistencia a la apoptosis [24].

## 1.6. Aproximaciones bioinformáticas para el estudio del perfil molecular del HCC

El perfil molecular del HCC ha sido caracterizado durante varios años utilizando técnicas de secuenciación y enfoques multi-ómicos. Las investigaciones en esta área se han basado, en gran medida, en análisis bioinformáticos para detallar las características moleculares del HCC. Las técnicas bioinformáticas han permitido no sólo la caracterización de mutaciones puntuales y cambios en la expresión, sino también la descripción de ANC. Algunos estudios, como el realizado por Gao y colaboradores, han analizado datos genómicos y transcriptómicos públicos de TCGA (The Cancer Genome Atlas Program) e ICGC (The International Cancer Genome Consortium). Estas aproximaciones facilitan la identificación de genes y redes genéticas involucradas en la patogénesis del HCC, ofreciendo una oportunidad para identificar marcadores moleculares [25].

Utilizando algoritmos de análisis de datos, se han identificado regiones cromosómicas que sufren amplificaciones y/o deleciones en pacientes con HCC, como en *chr1q* y *chr8p*, demostrando una correlación positiva entre la ANC y el estadio del tumor [21]. En otro estudio de Nakeep, mediante métodos *in silico*, se reportaron diversos marcadores moleculares y genéticos, donde los genes identificados están implicados en la proliferación, apoptosis, autofagia y respuesta inflamatoria. También se descubrió que el HCC asociado con VHB está relacionado con la mutación del gen *TP53* y del promotor *TERT* [7, 25].

Investigaciones más avanzadas que integran datos multi-ómicos, incluyendo datos genómicos (provenientes del TCGA), transcriptómicos y proteómicos se llevaron a cabo utilizando métodos bioinformáticos para identificar genes asociados con HCC [26]. Dicho estudio reconoció 8 genes centrales (*BUB1*, *BUB1B*, *CCNA2*, *CCNB1*, *CDC20*, *CDK1*,

*MAD2L1* y *CCNB2*) que están sobreexpresados y se relacionan con una supervivencia deficiente en pacientes con HCC. Estos genes están implicados en la regulación de la proliferación celular, la división celular y la formación de óxido nítrico. Por lo tanto, su sobreexpresión contribuye al crecimiento de células hepáticas transformadas.

Otro estudio realizado por Hu y colaboradores [27], utilizando datos de perfil de expresión génica de la base de datos Gene Expression Omnibus (GEO), identificó cinco genes clave incluyendo *RRM2*, *MAD2L1*, *MELK*, *NCAPG* y *ASPM*. Estos genes se asocian con una baja supervivencia en HCC y están involucrados en la proliferación, migración, invasión e inhibición de la apoptosis de células hepáticas transformadas. Además, Bassaganyas y colaboradores [28] mediante análisis de datos de matrices de SNP (polimorfismos de nucleótido único) demostraron que no había una relación entre la ANC de los genes del sistema inmunológico y las variables clínico-patológicas analizadas (VHB/VHC).

Por tanto, la accesibilidad a datos ómicos en recursos públicos como TCGA, ICGC, GEO junto con la aplicación de métodos bioinformáticos ha facilitado la disección biológica del HCC. Esto ha permitido describir mecanismos moleculares y bioquímicos específicos de la neoplasia haciendo posible el descubrimiento de potenciales marcadores.

## **1.7. Base de datos utilizada en el estudio**

En este estudio, los datos fueron obtenidos a través de Xena Browser, una plataforma de la Universidad de California, Santa Cruz (UCSC), que permite acceder y analizar datos genéticos y clínicos de acceso público. Específicamente, se utilizaron datos del proyecto TCGA, en particular de la cohorte GDC TCGA-LIHC (Liver Hepatocellular Carcinoma), que recopila información detallada de pacientes con HCC.

Los datos incluyen:

- Información genómica: mutaciones, ANCs
  - Regulación molecular: perfiles de expresión génica y metilación del ADN.

- Características clínicas y demográficas: edad, etnia, género, estadio tumoral, sobrevida, tratamiento.

El acceso a estos datos ha permitido aplicar análisis bioinformáticos avanzados para identificar biomarcadores y alteraciones genómicas clave en la progresión del HCC, contribuyendo al estudio de su biología y posibles estrategias terapéuticas.

## II. Planteamiento del problema

La inflamación es un factor crucial en el desarrollo del HCC, especialmente en la carcinogénesis asociada a infecciones crónicas por VHB/VHC. Estudios recientes, mediante un análisis integrador de genómica y transcriptómica, reportaron el incremento en la expresión de un conjunto de 200 genes relacionados con la respuesta inflamatoria en el HCC asociado a virus, tales como *CCL20*, *IL1B*, *CCL2* y entre otros. Se encontró que la sobreexpresión de dichos genes se correlaciona con una disminución en el tiempo de supervivencia [24], sugiriendo su potencial uso como marcadores pronósticos del HCC.

A pesar de estos hallazgos, no se ha investigado la ANC de estos genes (*CCL20*, *IL1B*, *CCL2* y entre otros) y su relación con la sobrevida global. La ANC, junto con los cambios en la expresión génica y las mutaciones, constituyen mecanismos comunes de alteración genética en el cáncer. En este contexto, la amplificación de un gen generalmente se asocia con su sobreexpresión, mientras que su delección suele llevar a la subexpresión [29]. Dado que hasta un 14% de los pacientes con HCC presenta ANC en su ADN [18], es fundamental estudiar la ANC de los genes reguladores de la inflamación.

El estudio de la ANC y su relación con el pronóstico es particularmente relevante en el HCC, debido a la notable reducción en la tasa de supervivencia relativa a 5 años según las distintas etapas del cáncer, llegando a ser de sólo 2% en etapas avanzadas [5]. A esto se le suma que la mayoría de los pacientes reciben un diagnóstico de HCC en una etapa avanzada

(metástasis), limitando su tratamiento únicamente a cuidados paliativos. Finalmente, un 20% de los individuos con HCC no sobreviven más de un año tras ser diagnosticados [30].

Por lo tanto, es de interés determinar si los genes de inflamación presentan ANC y si así fuera, si ello está influenciado por la presencia del VHB/VHC al momento del diagnóstico. Además, identificar los genes con ANC y su relación con la sobrevida puede revelar mecanismos que facilitan la progresión del HCC, pudiendo en el futuro ser evaluados como marcadores pronósticos que mejoren el manejo y la esperanza de vida de los pacientes.

A pesar de su importancia, las investigaciones genómicas del cáncer son complejas y requieren un gran número de participantes y tecnologías de secuenciación avanzadas, lo cual puede ser una limitación. Para superar esta dificultad, el presente estudio utilizará bases de datos públicas de genómica del cáncer y herramientas de análisis bioinformático. El acceso a estos datos y metadatos (información demográfica y clínica) es eficiente, y las herramientas bioinformáticas empleadas están diseñadas para garantizar la integridad y consistencia de los datos.

### **III. Pregunta de Investigación**

¿Existe alteración del número de copias de genes reguladores de inflamación en la cohorte internacional GDC TCGA - LIHC? Y si ese el caso, ¿la alteración se asocia con la sobrevida global y con la infección por VHB/VHC detectada al momento del diagnóstico?

### **IV. Objetivos**

#### **4.1. Objetivo general:**

Identificar la presencia de la ANC de genes reguladores de inflamación y su asociación con la OS e infección por VHB/VHC, en pacientes de la cohorte internacional GDC TCGA - LIHC, mediante un análisis secundario de datos genómicos.

#### **4.2. Objetivos específicos:**

- Identificar los genes reguladores de inflamación que presentan ANC en HCC.
- Relacionar la ANC con el estado viral (VHB y VHC) determinado al momento del diagnóstico del HCC.
- Asociar la ANC con la OS en pacientes con HCC agrupados por el estado viral (VHB y VHC).

### **V. Materiales y Métodos**

#### **5.1. Diseño de estudio**

Este estudio es de tipo observacional, retrospectivo y longitudinal. Se realizó un análisis secundario de datos genómicos y metadata de pacientes con HCC (reclutados en USA, principalmente y no hay datos peruanos), cuyo estado viral se conoce al momento del diagnóstico del cáncer. Dichos datos están disponibles en recursos públicos (ver sección 6.4).

#### **5.2. Población de estudio**

Son pacientes diagnosticados con HCC que forman parte de la cohorte de 469 participantes del proyecto GDC TCGA- LIHC [Descargado en Xena Browser, ver Anexo1]. Para elegir a los participantes del presente estudio, se aplicaron los siguientes criterios de elegibilidad:

Criterios de inclusión:

- Diagnóstico de HC

- Tener reportado el estatus de infección por VHB y VHC

Criterios de exclusión:

- Metadata no reportado: Estadio de tumor no reportado, OS, grupo étnico
- Pacientes cuyo ANC no esté disponible
- Pacientes con coinfección por VHB y VHC
- Pacientes con recaída: “Recurrent tumor”

Los criterios de inclusión se establecieron para seleccionar una cohorte homogénea de pacientes con HCC y estado viral registrado, garantizando la validez de los análisis genómicos. Se excluyeron casos con datos incompletos, coinfección o tumores recurrentes, ya que estos últimos pueden presentar alteraciones genéticas influenciadas por tratamientos previos y evolución tumoral. Esto podría introducir variabilidad y afectar la precisión del análisis. Al centrarse en tumores primarios, se asegura una evaluación más fiable de los mecanismos genéticos iniciales del HCC.

**Ecuación 1. Fórmula para el cálculo del tamaño muestral:**

$$n = \frac{N\sigma^2Z^2}{(N-1)e^2 + \sigma^2Z^2}$$

$$n = \frac{469(0,5)^2(1,96)^2}{(469-1)0,05^2 + 0,5^21,96^2}$$

Donde:

- Universo (N): 469
- Z-score: 1,96
- Proporción esperada (e): 0,5
- Margen de error (σ): 0,05
- Nivel de confianza: 95 %
- Muestra (n): 211,42

El número de registros analizados en la tesis, que corresponde a 210 participantes y tras aplicar los criterios de elegibilidad, coincide con el número estimado mediante el cálculo del tamaño muestral.

### 5.3. Operacionalización de variables

Las variables de estudio se presentan en la tabla 1.

**Tabla 1.** Operacionalización de variables de estudio

Variable	Definición práctica de la variable	Identificar tipo de variable	Identificar la escala de medición	Medición
Estado viral	Infección viral por Virus de la Hepatitis B (VHB) y Hepatitis C (VHC).	V. independiente (explicatoria), cualitativa dicotómica	Nominal <b>VHB+:</b> Individuo positivo para anticuerpos anti-VHB. <b>VHC+:</b> Individuo positivo para anticuerpos anti-VHC. <b>No infección por VHB (VHB-):</b> Individuo negativo para anticuerpos anti-VHB. <b>No infección por VHC (VHC-):</b> Individuo negativo para anticuerpos anti-VHC.	Prueba de antígeno

ANC	Alteración del número de copias en segmentos genómicos específicos ya sea mediante amplificaciones y/o deleciones.	V. dependiente (respuesta), cualitativo politómica	Nominal - <b>Amplificación del gen:</b> aumento en el número de copias del gen - <b>Deleción del gen:</b> disminución en el número de copias del gen. - <b>Sin cambio:</b> el número de copias no cambia.	Secuenciación del genoma completo (WGS), Análisis de microarrays
Sobrevida global (OS)	Se representa como el tiempo (en meses o años) desde el inicio	V. dependiente (respuesta), cuantitativa	De razón	Seguimiento médico regular para registrar la

	del diagnóstico/ tratamiento hasta el fallecimiento.	Continua		fecha de diagnóstico, el inicio del tratamiento y la fecha de fallecimiento (si corresponde).
Tipo de muestra	Muestra de tejido hepático para ver la presencia de tumor o ausencia de ello.	Variable independiente (explicativa), cualitativa dicotómica.	Nominal - Tumor - Normal	Dato clínico
Genes de inflamación	Genes que participan en la regulación o mediación de procesos inflamatorios.	V. independiente (explicativa), cuantitativa discreta	Nominal	PCR, secuenciación del ADN
Estadio de tumor	Representa el grado de avance y extensión del tumor en el cuerpo de un paciente.	V. independiente (explicativa) cualitativa politómica	Ordinal - Estadio I (leve) - Estadio II - Estadio III - Estadio IV (más grave)	Pruebas de imagen como resonancia magnética o tomografía computarizada

Etnia	Características físicas distintivas de una población de un mismo origen geográfico.	V. demográficas (descriptoras), cualitativa politómica	Nominal	Dato clínico
Edad	Tiempo en meses o años que ha vivido un individuo desde su nacimiento.	V. demográficas (descriptoras), cuantitativa continua	De razón	Dato clínico
Sexo	Clasificación de un individuo en masculino o femenino.	V. demográficas (descriptoras), cualitativa dicotómica	Nominal	Dato clínico

## 5.4. Recursos bioinformáticos

La presente tesis hizo uso de varias herramientas bioinformáticas, las cuales se presentan en la Tabla 2

**Tabla 2.** Recursos bioinformáticos

Nombre de la plataforma	Definición	Aplicación	Website
AmiGO 2	AmiGO 2 es una herramienta bioinformática donde se puede explorar, buscar y visualizar datos relacionados con ontologías de genes y términos de la Gene Ontology (GO). AmiGO 2 nos permite analizar conjuntos de genes para determinar las funciones biológicas, procesos y componentes celulares en los que están implicados.	Consultando esta herramienta se obtuvo la ontología de los genes reguladores de inflamación para este respectivo estudio.	<a href="https://soybase.org/amigo/amigo/landing">https://soybase.org/amigo/amigo/landing</a> Versión de AmiGO 2 versión: 2.5.17
Uniprot	Es una base de datos bioinformática que ofrece información detallada sobre secuencias y funciones de proteínas.	Permitió identificar la ontología de los genes reguladores de inflamación	<a href="https://www.uniprot.org/">UniProt</a>

EnrichR	Es una plataforma web bioinformática que facilita el análisis de enriquecimiento de conjuntos de genes.	Se utilizó para agrupar a los genes de acuerdo a sus rutas bioquímicas.	<a href="http://maayanlab.cloud">Enrichr (maayanlab.cloud)</a>
---------	---	---	--

UCS Xena Browser	Xena Browser es una plataforma que permite el análisis y visualización interactiva de datos genómicos y clínicos. Se pueden integrar conjuntos de datos de diversas fuentes y analizarlos de manera conjunta. Asimismo, es posible visualizar datos de expresión génica, mutaciones y metilación del ADN.	Sirvió para descargar la metadata. También introducir la lista de genes reguladores de inflamación con deleciones o amplificaciones y correr la Kaplan Meier.	<a href="http://xenabrowser.net">UCSC Xena (xenabrowser.net)</a>
cBioPortal for Cancer Genomics	cBioPortal es una plataforma centrada en la genómica del cáncer. Es posible visualizar datos de secuenciación, amplificación/delección de genes, mutaciones y expresión génica en conjuntos de muestras de pacientes con cáncer.	En esta plataforma se descargaron los genes reguladores de inflamación con deleciones o amplificaciones	<a href="http://cBioPortal.org">cBioPortal for Cancer Genomics</a>

Genome Data Viewer de NCBI	Esta plataforma brinda acceso a secuencias de genomas completos de diversas especies y te permite explorar la organización genómica, identificar genes, regiones donde	Nos permitió identificar las regiones en el cromosoma donde se ubican los genes reguladores de	<a href="https://www.ncbi.nlm.nih.gov/genome/gdv/">https://www.ncbi.nlm.nih.gov/genome/gdv/</a>
	se localizan estos genes en el cromosoma y variantes genéticas.	inflamación identificados en este estudio.	

## 5.5. Obtención de datos genómicos y metadata

### 5.5.1. Construcción de base de datos de la población

Todos los datos genómicos y la metadata del estudio se obtuvieron de la cohorte GDC TCGA-LIHC a través de la plataforma Xena Browser (Anexo 1). La información viral fue reportada en los estudios de Gao et al. y Ally et al. [24,19], quienes utilizaron cohortes de TCGA-LIHC. En ambos estudios, la detección de la información viral se realizó mediante pruebas de secuenciación de mRNA.

### 5.5.2. Lista de genes de interés

Los genes de inflamación que fueron de interés para este estudio se obtuvieron de la publicación de Gao y colaboradores [24], siendo 200 genes en total (Ver Anexo 2).

### 5.5.3. Obtención de la ontología de los genes de inflamación

En la plataforma de Amigo2, se eligió la opción avanzada de PANTHER y posteriormente se introdujo la lista de los 200 genes de inflamación (Anexo 2), luego se seleccionó el organismo (en este caso *Homo sapiens*) y por último enviamos.

**AmiGO 2**  
More information on quick search [?](#)

Quick search   
Search

**PubMed Search**

PubMed ID search takes you to summary information and all GO annotations for the PubMed ID.

E.g. PMID:123456

**Search Templates**

Use predefined templates to explore Gene Ontology data.

**Advanced Search**

Interactively search the Gene Ontology data for annotations, gene products, and terms using a powerful search syntax and filters.

**Browse the Ontology**

Use the drill-down browser to view the ontology structure with annotation counts.

**Term Enrichment Service**

Your genes here...

biological process

Homo sapiens

Powered by PANTHER

**Statistics**

View the most recent statistics about the Gene Ontology data in AmiGO.

**And Much More...**

Many more tools are available from the software list, such as alternate searching modes, Visualize, non-JavaScript pages.

**Help Tips**

**Steps:**

- > 1. Select list and list type to analyze
- > 2. Select Organism
- > 3. Select operation

[Using enhancer data](#)

**1. Enter ids and or select file for batch upload. Else enter ids or select file or list from workspace for comparing to a reference list.**

Enter IDs:  **1. Introducir la lista de genes de interés** separate IDs by a space or comma

Upload IDs:  No se eligió ningún archivo

Please [login](#) to be able to select lists from your workspace.

Select List Type:

- ID List
- Previously exported text search results
- Workspace list
- PANTHER Generic Mapping
- ID's from Reference Proteome Genome

Organism for id list:

VCF File Flanking region:   Search Enhancer Data

---

**2. Select organism.**

Mus musculus  
Rattus norvegicus  
Gallus gallus  
Danio rerio

---

**3. Select Analysis.**

- Functional classification viewed in gene list
- Functional classification viewed in graphic charts  Bar chart  Pie chart
- Statistical overrepresentation test
- Statistical enrichment test

**3. submit**

Para visualizar la ontología se personalizó la vista, haciendo click en “Customize Gene list” y del cuadro que se genera se añadieron los tres componentes de la ontología (componente celular, función molecular y proceso celular), luego se envió y finalmente se cerró. Finalmente para exportar la tabla en excel en la opción de “Send list to” se eligió “File”.

20

**4. Añadir los componentes de la Ontología**

PANTHER GENE LIST [Customize Gene list](#) [Click to view Enhancer Data](#)

Convert List to: [-Select-] Send list to: [-Select-]

Display: [30] items per page [Refresh](#) [Search](#)

Hits 1-30 of 200 [ page: (1) 2 3 4 5 6 7 ] Number of mapped: Text

5.

Gene ID	Mapped IDs	Gene Name Gene Symbol	PANTHER Family/Subfamily
<input type="checkbox"/> 1. <a href="#">HUMAN HGNC=348 UniProtKB=P35869</a>	AHR	Aryl hydrocarbon receptor AHR <a href="#">PTN002472519</a> <a href="#">orthologs</a>	<a href="#">ARYL HYDROCARBON RECEPTOR (PTHR10649:SF9)</a>
<input type="checkbox"/> 2. <a href="#">HUMAN HGNC=23992 UniProtKB=Q96DX8</a>	RTP4	Receptor-transporting protein 4 RTP4 <a href="#">PTN002494639</a> <a href="#">orthologs</a>	<a href="#">RECEPTOR-TRANSPORTING PROTEIN 4 (PTHR14402:SF8)</a>
<input type="checkbox"/> 3. <a href="#">HUMAN HGNC=8506 UniProtKB=P13725</a>	OSM	Oncostatin-M OSM <a href="#">PTN002494451</a> <a href="#">orthologs</a>	<a href="#">ONCOSTATIN-M (PTHR14261:SF0)</a>
<input type="checkbox"/> 4. <a href="#">HUMAN HGNC=16647 UniProtKB=O00574</a>	CXCR6	C-X-C chemokine receptor type 6 CXCR6 <a href="#">PTN002470989</a> <a href="#">orthologs</a>	<a href="#">C-X-C CHEMOKINE RECEPTOR TYPE 6 (PTHR10489:SF705)</a>
<input type="checkbox"/> 5. <a href="#">HUMAN HGNC=9669 UniProtKB=P23469</a>	PTPRE	Receptor-type tyrosine-protein phosphatase epsilon	<a href="#">RECEPTOR-TYPE TYROSINE-PROTEIN PHOSPHATASE EPSILON (PTHR19134:SF499)</a>

Para completar la ontología de los genes que en PANTHER no lograron llenar los 3 componentes, se utilizó UniProt. Una vez dentro de Uniprot, en la opción “Lista” se introdujeron los códigos de cada gen que son reconocidos por Uniprot (en la tabla obtenida de PANTHER figura dicho código), luego se hizo click en “Maps ID”.

UniProt BLAST Align Peptide search ID mapping SPARQL Release 2024\_04 | Statistics Help

## Find your protein

UniProtKB  Advanced | **List** | Search

Examples: [insulin](#), [APP](#), [Human](#), [P05067](#), [organism\\_id:9606](#)

UniProt is the world's leading high-quality, comprehensive and freely accessible resource of protein sequence and functional information. [Cite UniProt](#)

**Proteins**  
UniProt Knowledgebase

**Species**  
Proteomes

**Protein Clusters**  
UniRef

**Sequence archive**  
UniParc

UniProt BLAST Align Peptide search ID mapping SPARQL UniProtKB Advanced | List Search

## Retrieve/ID mapping

Enter one or more IDs (100,000 max). You may also load from a text file. Separate IDs by whitespace (space, tab, newline) or commas. This service can also be used programmatically. Alternatively, the underlying data can be downloaded.

P31146 P02558 ALBU\_HUMAN EFTU\_ECOLI

**1. Introducir los códigos de UniProt de los genes de interés**

From database To database

UniProtKB AC/ID UniProtKB

Name your ID Mapping job

"my job title"

Restablecer Map IDs

2

Feedback Help

Una vez generada la tabla, se personalizó para verificar la ontología haciendo click en “Customize columns”, en la fila “Gene Ontology” se añadió los 3 componentes y se guardó. Para descargar la tabla, se seleccionaron todos los genes y se eligió el formato “Excel” (Anexo 2.1, Anexo 2.2, Anexo 2.3).

#### 5.5.4. Localización de los genes reguladores de inflamación en los diferentes cromosomas

En vista de que se ha demostrado la alta frecuencia de ANC de los cromosomas cerca de la integración del ADN viral del VHB en las células del hospedero [18] se identificó las regiones del cromosoma donde se alojan los genes reguladores de inflamación (se tomó en cuenta los 200 genes), para evaluar si son regiones donde ya se observó integración viral. Para ello se utilizó la herramienta Genome Data Viewer de NCBI (<https://www.ncbi.nlm.nih.gov/gdv/>):

- >1. En la página principal, en el buscador se introdujo el gen regulador de inflamación de nuestro interés.
- >2. Se generó una ventana con una fila en el extremo izquierdo donde se visualizó el cromosoma y las coordenadas del genoma en donde se ubica el gen introducido.

Asimismo, en el lado derecho es posible ver en qué brazo del cromosoma se ubica dicho gen.

Para realizar este proceso se ingresaron los genes de uno en uno y se construyó la BD en el excel de forma manual (Anexo 3)

## Genome Data Viewer

GDV supports the exploration and analysis of [NCBI-annotated](#) and selected non-NCBI annotated eukaryotic genome assemblies. Currently, assemblies from over 3160 organisms are available.

The screenshot shows the Genome Data Viewer interface. On the left, there is a 'Switch view' button and a 'Search organisms' box containing 'Homo sapiens (human)'. Below this is a tree view of organisms with a red arrow pointing to the 'human' node, labeled 'Introducir el gen de interés'. On the right, the 'Homo sapiens (human)' panel is visible, featuring a search bar for the genome, a dropdown for the assembly (GRCh38.p14), and buttons for 'Browse genome', 'Compare genomes', and 'Assembly details'.

## Genome Data Viewer

1. Introducir los genes de interés

The screenshot shows the Genome Data Viewer interface with the search results for the ABCA1 gene. A red arrow points to the search bar where 'ABCA1' has been entered. Below the search bar, a table lists the genes found. The first row, 'ABCA1', is highlighted with a red box. The table has columns for 'Name' and 'Location'. The 'Location' for ABCA1 is 'Chr9: 104,781,006 - 104,928,155'. To the right of the table, there is a genomic track showing the location of the gene on chromosome 9, with a red box highlighting the region '104,780 K - 104,820 K'. Below the track, there are various genomic tracks including 'Live RefSNPs, dbSNP b156 v2', 'RNA-seq exon coverage, aggregate (filtered), NCBI Homo sapiens Annotation Release 110 - log', and 'RNA-seq intron-spanning reads, aggregate (filtered), NCBI Homo sapiens Annotation Release 11'. A second red arrow points to the genomic track with the text '2. Identificar la región donde se ubica el gen'.

Name	Location
ABCA1	Chr9: 104,781,006 - 104,928,155
CAV1	Chr7: 116,525,009 - 116,561,185
APOE	Chr19: 44,905,796 - 44,909,393
TNF	Chr6: 31,575,565 - 31,578,336
TNF	NT_113891.3: 3,052.9K - 3,055.6K
TNF	NT_167244.2: 2,909.1K - 2,910.9K
TNF	NT_167245.2: 2,823.3K - 2,826.1K
TNF	NT_167246.2: 2,880.6K - 2,883.4K

### 5.5.5. Obtención de la ANC en la cohorte GDC TCGA-LIHC

Para cumplir con el objetivo de identificar los genes reguladores de inflamación que presentan ANC, se utilizó la plataforma de CBioPortal, en donde se buscó la Cohorte GDC TCGA-LIHC y se inició a explorar dicho estudio.

En la parte superior del lado izquierdo se seleccionó la opción “Custom selection” y se introdujo el ID de todos los participantes que se ilustraron en el Anexo 4 en la primera fila del documento. Asimismo, en ese mismo lado se introdujo la lista de genes de inflamación a evaluar (200 genes en total, extraídos de la primera tabla del Anexo 2). Para descargar los datos, se hizo click en el icono de descarga que se encuentra en la parte superior de la página y para poder visualizarlo se abrió con excel (Anexo 5).

The screenshot shows the cBioPortal interface for selecting studies. At the top, there is a navigation bar with links for Data Sets, Web API, Tutorials/Webinars, FAQ, News, Visualize Your Data, About, and cBioPortal Installations. Below this is a green banner with a newsletter subscription link. The main content area is titled 'Select Studies for Visualization & Analysis' and shows 462 studies available (278369 samples). A dropdown menu for 'Data type' is highlighted with a red arrow. Below this, there is a search bar with the text 'Seleccionar la cohorte de interés'. The interface is divided into two columns: a left column with study categories and a right column with a list of specific studies. The categories include PanCancer Studies (10), Pediatric Cancer Studies (14), Immunogenomic Studies (8), Cell lines (3), PreCancerous Studies (1), Adrenal Gland (3), Ampulla of Vater (1), Biliary Tract (16), and Bladder/Urinary Tract (22). The list of studies includes MSK-IMPACT Clinical Sequencing Cohort (10945 samples), Metastatic Solid Cancers (500 samples), MSS Mixed Solid Tumors (249 samples), SUMMIT - Neratinib Basket Study (141 samples), TMB and Immunotherapy (1661 samples), Tumors with TRK fusions (106 samples), Cancer Therapy and Clonal Hematopoiesis (24146 samples), China Pan-cancer (10194 samples), Pan-cancer analysis of whole genomes (2922 samples), and MSK MetTropism (25775 samples). A red arrow points to the 'Explore Selected Studies' button at the bottom right.

The screenshot shows the cBioPortal interface for Liver Hepatocellular Carcinoma. The top navigation bar includes links for Data Sets, Web API, Tutorials/Webinars, FAQ, News, Visualize Your Data, About, and cBioPortal Installations. A green banner at the top contains a newsletter subscription link. Below this, the main header displays the study name and a search bar. A red arrow points to the search bar with the text '4. Introducir la lista de los genes'. Below the search bar, there are buttons for 'Custom Selection', 'Charts', and 'Groups'. A dropdown menu is open, showing 'By sample ID' and 'By patient ID' options. A red arrow points to the 'By patient ID' option with the text '3. Introducir la lista de los IDs de los participantes'. The main content area is divided into several panels: 'Cancer Type Detailed' showing Hepatocellular Carcinoma (369, 99.2%) and Fibrolamellar Carcinoma (3, 0.8%); 'Molecular Profile' with a list of genomic features and their counts; 'KM Plot: Disease-specific (months)' and 'KM Plot: Progression Free (months)'; 'Mutation Count vs Fraction Genome Altered'; 'Mutated Genes (366 profiled samples)' with a table showing TP53 (113 mutations, 30.1% frequency); and 'Structural Variant Genes (366 profiled samples)' with a table showing PRKACA (7 SVs, 1.9% frequency).

### 5.5.6. Incorporación de la metadata y ANC

Para incorporar el ANC obtenido en la sección anterior a la base de datos de pacientes construida (Anexo 4, que contiene la información clínica y demográfica de la cohorte), se utilizó R-Studio. Este lenguaje de programación cuenta con diversos paquetes para análisis estadísticos, para realizar esta tarea se usó el paquete de “dplyr” que tiene incorporado la función de “filter” y “merge”. Este último, fusiona base de datos por columnas o filas comunes [31], por ello en los anexos 4 y 5 (archivos utilizados para esta actividad) se colocó el ID del paciente en la primera fila. A continuación, se detallan los pasos seguidos:

Primero se instaló el paquete y luego se cargó:

```
install.packages("dplyr")
```

```
library(dplyr)
```

# Código para unir el archivo donde se visualiza la ANC y genes de inflamación por paciente y la información clínico-demográfica de los participantes.

```
merge(A4_BD_Pacientes, A5_ANC_Pacientes, by = "ID")
```

#Eliminar a los participantes que no tienen alteración en ningún gen (delección o amplificación)

```
TablaB <- TablaA%>%filter(!if_all(ABCA1:VIP,is.na))
```

```
#Eliminar los genes que no tengan ninguna alteración ya sea amplificación o delección  
TablaC <- TablaB[,colSums(is.na(TablaB))<nrow(TablaB)]  
write.xlsx(TablaC, "Anexo6.xlsx")
```

Con ello, se obtuvo la base de datos completa de los pacientes con sus datos demográficos, clínicos y genéticos.

### **5.5.7. Enriquecimiento funcional de los genes reguladores de inflamación con ANC**

Para agrupar a los genes que presentaron ANC (obtenidos en el apartado anterior que se muestran en el Anexo 6) según su ruta bioquímica, se utilizó la plataforma de Enrichr ([Enrichr \(maayanlab.cloud\)](https://maayanlab.cloud/enrichr/)). Se introdujo la lista de genes que presentaron delección o amplificación, una vez cargado en la sección “pathways” se seleccionó el KEGG 2021 (Kyoto Encyclopedia of Genes and Genomes). Luego se pudo visualizar la información en tabla o gráfico. Finalmente se exportó la tabla en formato excel haciendo click en “Export entries to table” (Anexo 7).

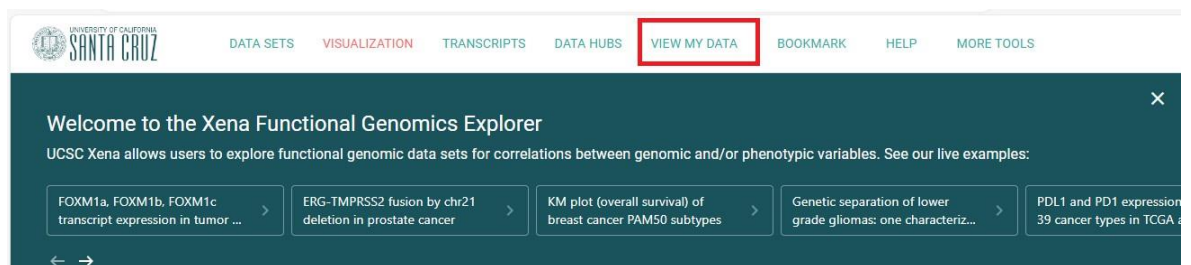
Las rutas bioquímicas o de señalización identificadas fueron validados por la ontología obtenida en la sección 6.5.3 y nos permitió reunir a los participantes que presentaron ANC solo en esos genes, es decir, se seleccionaron únicamente aquellos participantes que presentaron ANC (amplificación o delección) en los genes específicos dentro de cada ruta. Esto fue útil para correr el KM en Xena Browser.

### 5.5.8. Elaboración del plot de KM y ANC

Se trazaron curvas de Kaplan-Meier para evaluar la sobrevida global de los pacientes de estudio en relación con ciertos factores genéticos o clínicos de interés.

Para realizar dicho gráfico se siguieron estos pasos:

1. Se subió la matriz en excel (Anexo 8) que tiene el sample (para que en Xena lo reconozca), el ID del paciente, estado viral (VHB, VHC o Virus neg.), ANC (con cambios o sin cambios) y sobrevida global.
2. Se seleccionó el tipo de dato que presentaba la matriz (Fenotípico o genómico), en ese caso fue fenotípico.
3. Después se eligió la opción de donde estaba el sample ID si en la primera fila o columna.
4. Posteriormente, en la opción “There is other public data in Xena on these samples (e.g. TCGA) and want to connect to it” se seleccionó el estudio de “GDC TCGA Liver Cancer (LIHC)”.
5. A continuación, se procedió a importar y visualizar los datos.



My computer hub ● 1. LOAD DATA HELP

A Local Xena Hub is an application on your computer for loading and storing data.

We support most types of genomic and/or phenotypic/clinical/annotation data. This data can be your own or from another source, such as a publication.

Data on a Local Xena Hub can only be viewed or accessed by the same computer on which it is running.

## Loading data...

HELP

1 Select a Data File

2 Add Data Details

3 Load Data

Select Data File



HELP ON DATA FILE FORMATTING

CANCEL NEXT

## Loading data...

HELP

✓ Select a Data File

2 Add Data Details

3 Load Data

### File to Import: Matriz de BD.xlsx

HELP

Genomic data: numbers in a rectangle (e.g. expression)

Phenotypic data: categories or non-genomic in a rectangle (e.g. age, mutation status: 'wt' or 'mutant')



More data types...

#### File preview

Sample	"ID"	Status_viral	Status_ANC	Sobrevida_Global
TCGA-2Y-A9GT-01A	TCGA-2Y-A9GT	VHB	Con cambios	1624
TCGA-2Y-A9GU-01A	TCGA-2Y-A9GU	VHB	Con cambios	1939
TCGA-2Y-A9GV-01A	TCGA-2Y-A9GV	VHB	Con cambios	2532
TCGA-2Y-A9GX-01A	TCGA-2Y-A9GX	VHB	Sin cambios	2442

Show more...

## Loading data...

HELP

✓ Select a Data File

2 Add Data Details

3 Load Data

### File to Import: Matriz de BD.xlsx

The first column is sample IDs



The first row is sample IDs

Sample	"ID"	Status_viral	Status_ANC	Sobrevida_Global
TCGA-2Y-A9GT-01A	TCGA-2Y-A9GT	VHB	Con cambios	1624
TCGA-2Y-A9GU-01A	TCGA-2Y-A9GU	VHB	Con cambios	1939
TCGA-2Y-A9GV-01A	TCGA-2Y-A9GV	VHB	Con cambios	2532
TCGA-2Y-A9GX-01A	TCGA-2Y-A9GX	VHB	Sin cambios	2442

Show more...

1 Select a Data File — 2 Add Data Details — 3 Load Data

**File to Import: Matriz de BD.xlsx**

These are the first data on these samples.  
 I have loaded other data on these samples and want to connect to it.  
 There is other public data in Xena on these samples (e.g. TCGA) and want to connect to it. **4.**

Select Study  
GDC TCGA Liver Cancer (LIHC)

Sample	"ID"	Status_viral	Status_ANC	Sobrevida_Global
TCGA-2Y-A9GT-01A	TCGA-2Y-A9GT	VHB	Con cambios	1624
TCGA-2Y-A9GU-01A	TCGA-2Y-A9GU	VHB	Con cambios	1939
TCGA-2Y-A9GV-01A	TCGA-2Y-A9GV	VHB	Con cambios	2532
TCGA-2Y-A9GX-01A	TCGA-2Y-A9GX	VHB	Sin cambios	2442

[Show more...](#)

BACK CANCEL IMPORT **5.**

Es importante señalar que, para construir el plot de Kaplan-Meier (KM), se requieren dos variables clave: el tiempo hasta el evento de interés y una variable binaria que indique el estado de los pacientes. En este estudio, el tiempo hasta el evento se representa como la sobrevida global, que corresponde al período transcurrido desde un punto de inicio (como el diagnóstico o el inicio del tratamiento) hasta la ocurrencia de un evento, como la muerte o la recaída. La variable binaria, por su parte, refleja el estado del paciente, donde 1 indica que el paciente ha fallecido y 0 que sigue con vida [32]. Esta última variable se encuentra registrada en el estudio GDC TCGA-LIHC, por ello se vinculó con nuestra matriz en el paso 4.

6. Una vez importados los datos en la interfaz de Xena Browser, se seleccionó la opción "Phenotype". Luego, en la opción avanzada, se eligieron las variables de interés para realizar el análisis de KM.

7. Finalmente, se hizo click en la opción de "Kaplan Meier Plot" para visualizar el gráfico del KM.

**A** GDC TCGA Liver Cancer (LIHC)  
469 samples

**B** Select Your First Variable  
Add a phenotype (e.g. sample type, age).

**C** Select Your Second Variable

samples

50 samples

Select Data Type

Genomic

Phenotypic **6.**

Phenotype

statu

Basic Advanced

last\_known\_disease\_status.diagnoses

person\_neoplasm\_cancer\_status

Status\_ANC

Status\_viral

vital\_status.demographic

**cAMP**  
17 Samples

Type here or use dropper to select sample  
17 matching samples

ZOOM: None

[User Guide](#) - [Zoom Help](#) - [Tooltip Help](#)

A	B	C	D
Sample ID samples	cAMP.xlsx HCV_consensus	cAMP.xlsx anyinfection	cAMP.xlsx HBV_conse
samples	HCV_consensus	anyinfection	HBV_conse
TCGA-BW-A5NO	pos	pos	neg
TCGA-G3-A5SI	pos	pos	pos
TCGA-G3-A7M6	pos	pos	pos
TCGA-CC-A7IL	pos	pos	pos
TCGA-G3-A25X	pos	pos	pos
TCGA-QA-A7B7	pos	pos	pos
TCGA-UB-A7ME	pos	pos	pos

- Kaplan Meier Plot
- Chart & Statistics
- Reverse sort
- Download
- About
- Display
- Edit
- Remove

Se comparó la OS entre los participantes que presentaron ANC y aquellos que no lo hicieron. Además, se analizó la OS en el grupo de participantes con ANC,

diferenciando entre aquellos infectados con VHB, VHC y los que no presentaban ninguna infección.

Asimismo, para diferenciar la OS entre los participantes que presentaron amplificación o delección en ciertos genes, agrupados según sus rutas bioquímicas, se generaron matrices manualmente (ilustrados en el Anexo 9) a partir de la matriz general (Anexo 8). A diferencia de la matriz original, en estas nuevas matrices se indicó si el participantes tenía infección viral (pos) o no tenía (neg), así como el tipo de alteración en los genes, especificando si se trataba de delección (del), amplificación (amp) o ambas (amp/del).

> Se repitieron los pasos (1-7) para subir el resto de matrices que se muestran en el Anexo 8.

## **5.6. Análisis estadístico**

Todos los análisis estadísticos se realizaron en R (versión 4.3.2). En primer lugar, se efectuó un análisis estadístico descriptivo para resumir las características clínicas y demográficas de los pacientes, utilizando el paquete *compareGroups*. En este análisis, se describieron la media de la edad, así como la distribución por raza, sexo y estadio del tumor. Para la variable de supervivencia global, se empleó la mediana en lugar de la media, ya que los datos no seguían una distribución normal. Luego mediante un análisis estadístico inferencial se evaluó la asociación entre las variables mencionadas y el estado viral, con el objetivo de determinar si existían diferencias significativas entre los distintos grupos. Se consideró un valor de  $p < 0.05$  como indicador de significancia estadística.

Los códigos utilizados fueron los siguientes y la tabla utilizada fue el Anexo 4:

```

>library(compareGroups)
> tabla_descriptiva <- descrTable(Statuts_viral ~., Anexo4_BD_Pacientes)
> print(tabla_descriptiva)
> export2word(tabla_descriptiva, file="Tabla_Descriptiva.docx")

```

Por otro lado, se utilizó la prueba de Fisher en R Studio para evaluar si las diferencias entre la amplificación y la delección de cada gen eran estadísticamente significativas. Para ello, se creó manualmente una hoja de cálculo en Excel (Anexo 10) a partir del Anexo 6, en la cual se detalla el ID de cada participante, su estado viral y los genes con ANC.

En esta nueva tabla se registró el número de cambios en cada gen, es decir, cuántos individuos presentaban delección, amplificación o ningún cambio en los genes con ANC. Los participantes se agruparon según su estado viral: virus positivo (se contabilizó el número de participantes con virus que presentaban cambios en dicho gen) y virus negativo (se contabilizó el número de participantes sin virus que presentaban cambios en dicho gen). Tras recopilar los datos de todos los genes, estos se introdujeron de manera individual en R. El código utilizado para la prueba de Fisher en R fue:

```

> flyi = c(2,0,96,0,0,38) .... Estos valores varían según la cantidad de cambios
observados en cada gen. Los tres primeros valores corresponden a individuos con
infección viral: el 2 representa el número de individuos con delección en el gen, el 0
indica la cantidad de individuos con amplificación y el 96 corresponde a aquellos sin
ANC. Los últimos tres valores siguen el mismo esquema (delección, amplificación y
sin ANC), pero corresponden a individuos sin infección viral.
> fly = matrix(flyi, ncol=2, nrow=3)
> print(fly)
> fisher.test(fly)

```

Además, se utilizó la prueba de Chi-cuadrado en R para evaluar la relación entre la ANC de genes reguladores de inflamación y el estado viral (VHB, VHC). Los individuos se agruparon según el tipo de ANC (amplificación o delección) y, a su vez, se subdividieron en grupos con y sin infección viral. Se consideraron significativas las diferencias con un valor de  $p < 0.05$ . El código utilizado en R fue:

```
> valores = c (72,28) ...Los valores representan el porcentaje de amplificación o
delección (amp/del) en individuos con y sin infección viral.
## en el nombre de la lista puede ir delección o amplificación
> tablaf = matrix(valores, ncol=2, nrow=1,
  dimnames = list(c("amp"),
    c ("Neg","Pos")))
##con porcentajes por columnas
> round (prop.table(tablaf, 2)*100,2)
> chisq.test(tablaf, correct=F)
```

Por último, con un análisis de supervivencia de Kaplan–Meier se evaluó si la ANC en los genes reguladores de inflamación influía en la sobrevida global de los pacientes con HCC e infección viral como se detalló en la sección (6.5.7).

Para ello, se empleó la prueba log-rank disponible en Xena Browser, que permite comparar las distribuciones de supervivencia entre varios grupos expuestos a distintas condiciones. Esta prueba es no paramétrica que se basa en rangos y contrasta el número de muertes observadas en cada grupo con el número esperado bajo la hipótesis nula (se espera que la cantidad de fallecidos sea casi igual en ambos grupos que se está comparando). La estadística log-rank sigue aproximadamente una distribución Chi-cuadrado con grados de libertad equivalentes al número de grupos - 1 [32].

Ecuación 2. Fórmula para calcular la estadística log-rank:

$$\chi^2 = \sum \frac{(\sum O_{jt} - \sum E_{jt})^2}{\sum E_{jt}}$$

Donde;

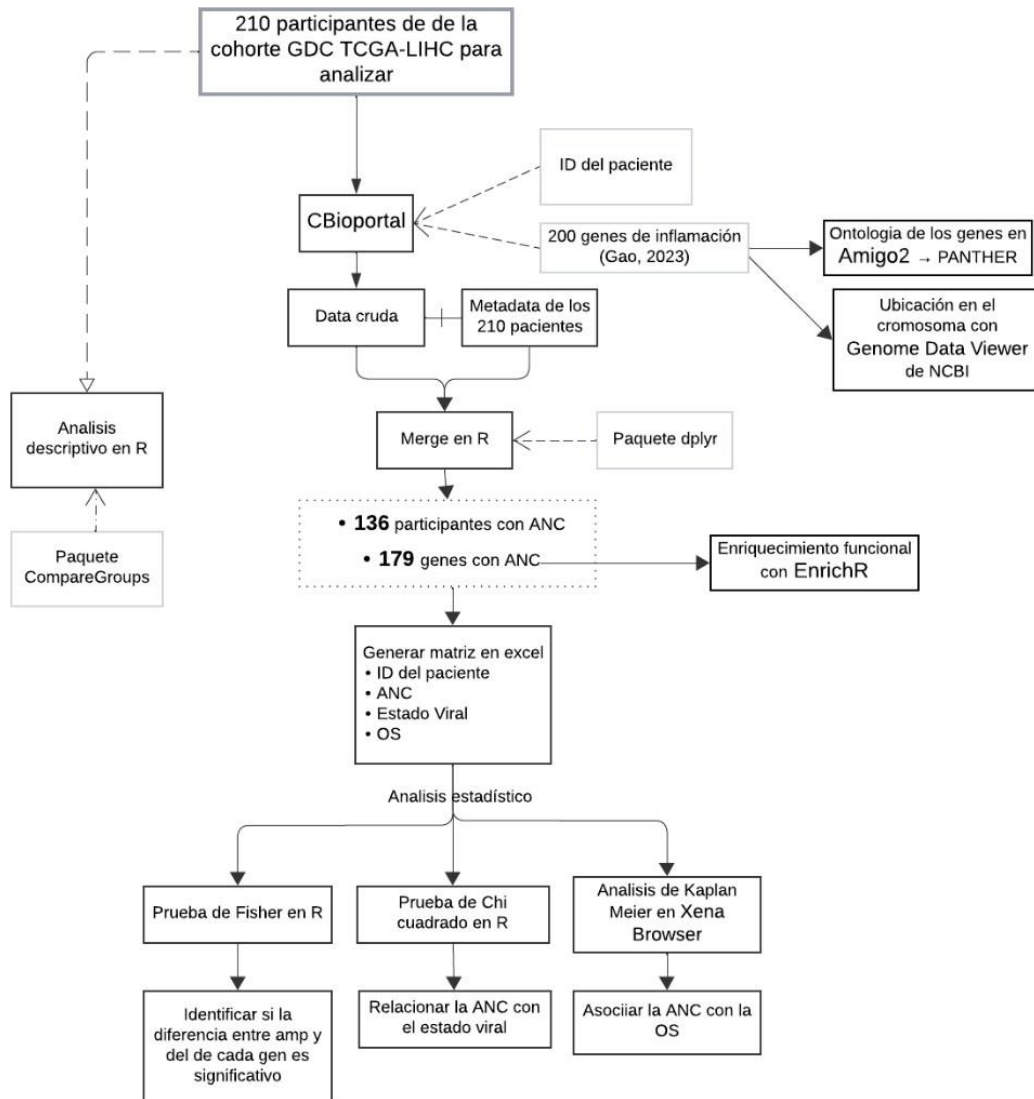
$\sum O_{jt}$  : suma del número observado de eventos en el grupo j a lo largo del tiempo

$\sum E_{jt}$ : suma del número esperado de eventos en el grupo j a lo largo del tiempo

j=1,2

Utilizando esta fórmula se puede calcular un valor de Chi cuadrado, a partir del cual es posible obtener el valor p, Se considera que existe significancia estadística cuando  $p < 0.05$ .

El resumen de las actividades realizadas en esta tesis se presenta en la figura 2.



S

Figura 1. Resumen de actividades realizadas en la tesis

### 5.7. Cuestiones éticas

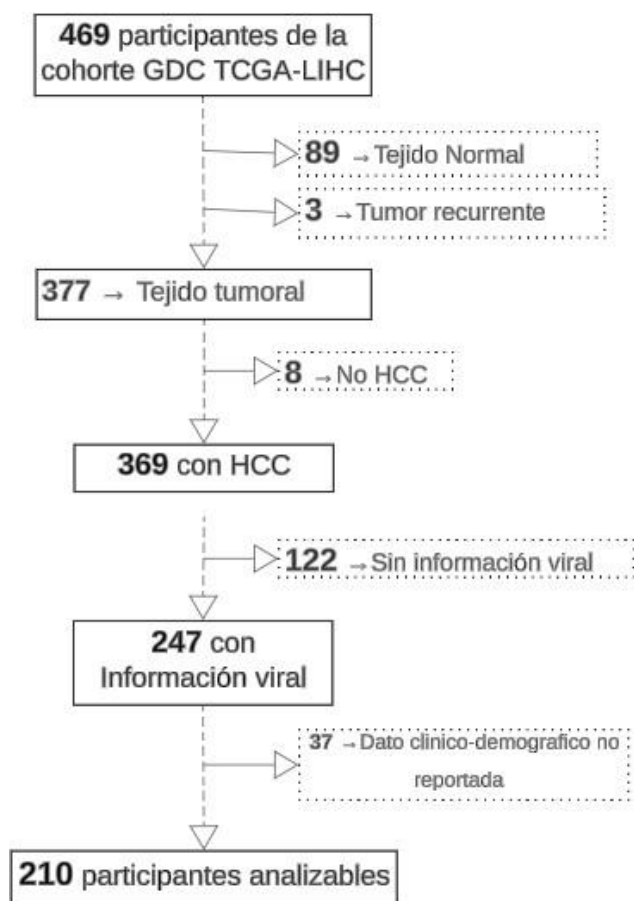
En este proyecto no se enrolaron participantes ni se hicieron pruebas *in vitro* o *in vivo*. Se utilizaron datos públicos de muestras anonimizadas de la cohorte GDC TCGA - LIHC, los cuales fueron reanalizados. Por esta razón, se obtuvo la exoneración de la aprobación del Comité de Ética de la UPCH, conforme a la carta SIDISI N° 213268, emitida y aprobada el 16 de febrero de 2024.

## **VI. Resultados**

### **6.1. Descripción clínica-demográfica de la cohorte GDC TCGA-LIHC**

De un total de 469 participantes que fueron extraídos de la plataforma bioinformática de Xena Browser, tras aplicar los criterios de elegibilidad (Sección 6.2), se excluyeron a 259 por no contar con información viral, presentar coinfección o no tener reportado el dato clínico demográfico (grupo étnico, estadio de tumor, sobrevida global), quedando así 210 participantes aptos para el estudio (ver figura 2, Anexo 4).

En este estudio, se consideró la etnicidad de los participantes, ya que esta se relaciona con la composición genética y cultural de la población. Los datos de etnicidad provienen del proyecto TCGA y fueron obtenidos a través de Xena Browser. Además, la residencia demográfica se define como el país en el que habita un individuo y donde puede estar expuesto a factores ambientales y condiciones socioeconómicas que influyen en su salud. Según lo reportado por Ally y colaboradores [19], los participantes residen en Estados Unidos, Vietnam, Corea del Sur y Canadá, lo que refleja una diversidad geográfica dentro de la cohorte estudiada.



**Figura 2.** Selección de Pacientes de la Cohorte GDC TCGA-LIHC

Después de seleccionar a los 210 participantes se observó que, en ambos grupos (VHB y VHC), más de la mitad de los participantes fueron hombres, resultando significativa la diferencia en la distribución por sexo entre los grupos ( $p < 0.009$ ) (Tabla 2). Por otro lado, el rango de edad de la población de estudio fue entre 17-90 años (Anexo 4), es muy probable encontrar adultos con infección viral (VHB o VHC) que sin infección ( $p < 0.05$ ). Por otro lado, la distribución de razas varía significativamente entre los grupos. Así, en el grupo VHB, hubo una mayor proporción de asiáticos y blancos (48% en ambos casos) en comparación con negros o afroamericanos, mientras que en el grupo VHC y en el grupo Sin infección, predominó una mayor proporción de blancos frente a asiáticos o negros.

En cuanto al estadio del tumor, aunque se observó una mayor proporción de individuos en estadios intermedios en los grupos de VHB, VHC y sin infección, en comparación con los estadios avanzados, no se encontró una diferencia significativa entre estos grupos. Por último, en cuanto a la sobrevida global, hay una diferencia significativa entre los 3 grupos por tipo de infección, encontrando que el grupo con VHC vive 300 días menos a diferencia del grupo con VHB o sin infección (Tabla 2).

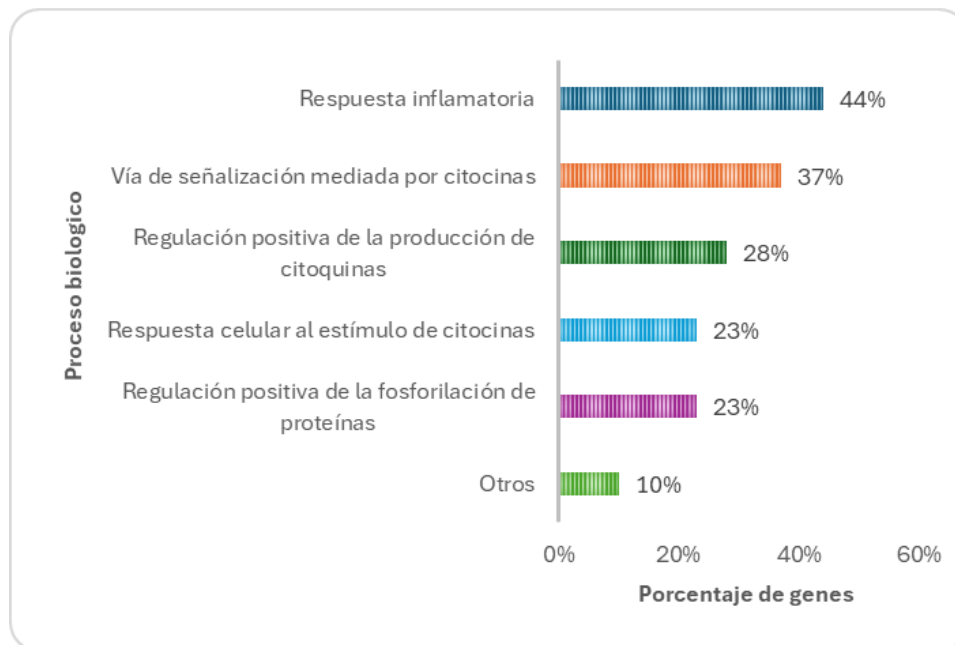
**Tabla 2.** Resumen de los datos clínicos-demográficos de los 210 participantes analizables.

	VHB <i>N=115</i>	VHC <i>N=49</i>	Sin virus <i>N=46</i>	p.overall
<b>Sexo:</b>				0.009
Femenino	40 (34.8%)	12 (24.5%)	25 (54.3%)	
Masculino	75 (65.2%)	37 (75.5%)	21 (45.7%)	
<b>Grupo etario:</b>				0.014
Joven (17-26años)	5 (4.35%)	1 (2.04%)	3 (6.52%)	
Adulto(27-59años)	56 (48.7%)	27 (55.1%)	34 (73.9%)	
Adulto mayor(>59años)	54 (47.0%)	21 (42.90%)	9 (19.6%)	
<b>Raza:</b>				<0.001
Asiatico	55 (47.8%)	11(22.4%)	0 (0.0%)	
Negro o afroamericano	5 (4.35%)	4 (8.16%)	1 (2.17%)	
Blanco	55 (47.8%)	34 (69.4%)	45 (97.8%)	
<b>Estadio de tumor:</b>				0.229
Estadio intermedio o avanzado (iii y iv)	71 (61.7%)	37 (75.5%)	31 (67.4%)	
Estadio temprano (i y ii)	44 (38.3%)	12 (24.5%)	15 (32.6%)	
<b>Sobrevida Global (días)</b>	834 (838)	529 (438)	855 (756)	0.040

## 6.2. Ontología y distribución cromosómica de los genes reguladores de inflamación

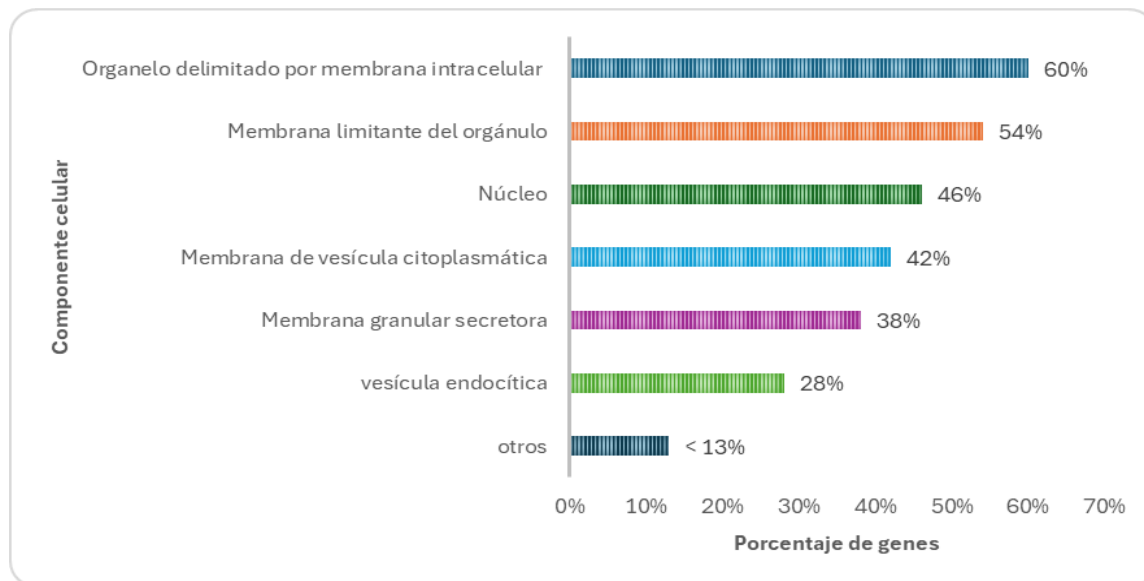
Se estudió la ontología para clasificar y entender las funciones biológicas, los procesos moleculares y las rutas en las que participan los genes reguladores de la inflamación. Por otro lado, también se exploró la distribución cromosómica para localizar las regiones en las que se encuentran los genes evaluados. Utilizando el Anexo 2 en formato Excel, se elaboraron gráficos de barras que agrupan los genes según su proceso biológico, componente celular y función molecular.

En relación al primer término de ontología (proceso biológico) de los genes de estudio, tal como se esperaba, se observó que casi la mitad de los genes (44%) interviene en la respuesta inflamatoria, incluyendo *PTGER4*; *CXCL6*; *CD40*; *CXCL9*; *CXCL8*; *TNFAIP6*; *PTGER2*; *C5AR1*; *FPR1*; *ADM*; *MEFV*; *CX3CL1*; *RELA*; *CCL7*; *IRAK2*; *CCL5*; *CCRL2*; *C3AR1*; *PROK2*; *OLR1*; *NLRP3*; *CCL2*; *CCR7*; *CD14*; *CCL17*; *CMKLR1*; *CCL24*; *PTGIR*; *CCL22*; *RIPK2*; *CCL20*; *IL18*; *CYBB*; *TACR1*; *SELE*; *NFKB1*; *TLR1*; *IL1A*; *CXCL10*; *CXCL11*; *IL6*; *IL1B*; *TLR3* y *TLR2* (Anexo 2.1, Figura 3).



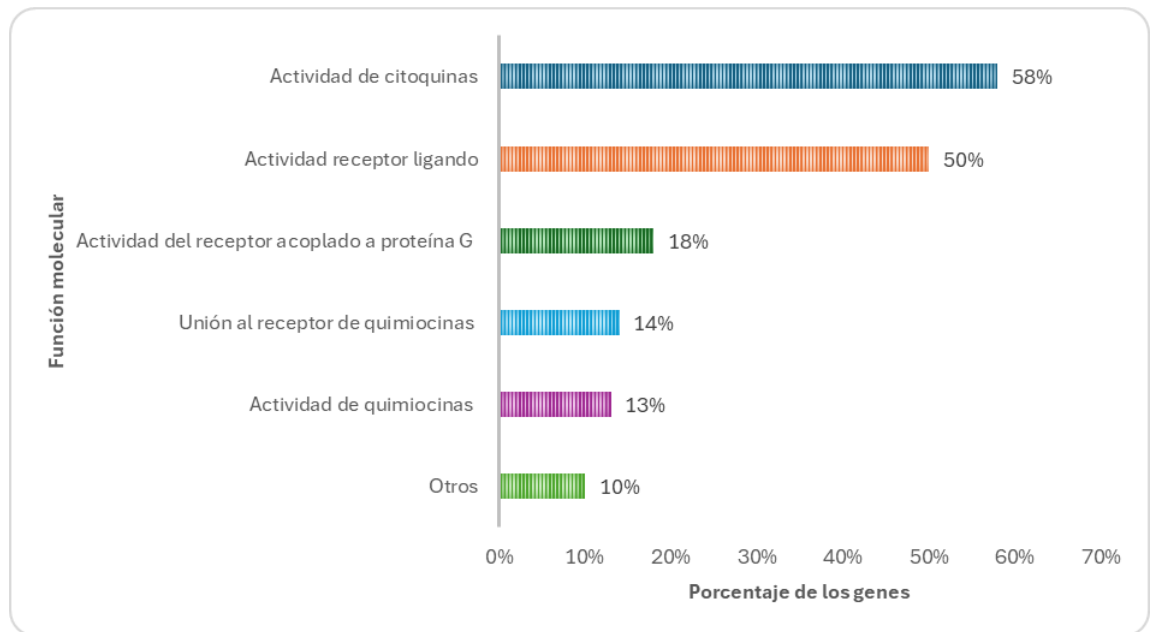
**Figura 3.** Proceso biológico de los 200 genes de inflamación.

Con respecto al componente celular, el 60% de genes (*BTG2; CDKN1A; ITGB3; AQP9; AHR; PSEN1; MEFV; HIF1A; RELA; IRAK2; MYC; LAMP3; PDE4B; OLR1; NLRP3; ABCA1; LYN; GCH1; SPHK1; SLC11A2; ATP2B1; NFKB1; BST2; NFKBIA; PTPRE; MMP14; KLF6; AXL; IRF1; IRF7*) forman parte organelas delimitado por membrana intracelular, seguido por membrana limitante del orgánulo, núcleo, membrana de vesícula citoplasmática, membrana granular secretora, vesícula endocítica y otros (Anexo 2.2, Figura 4).



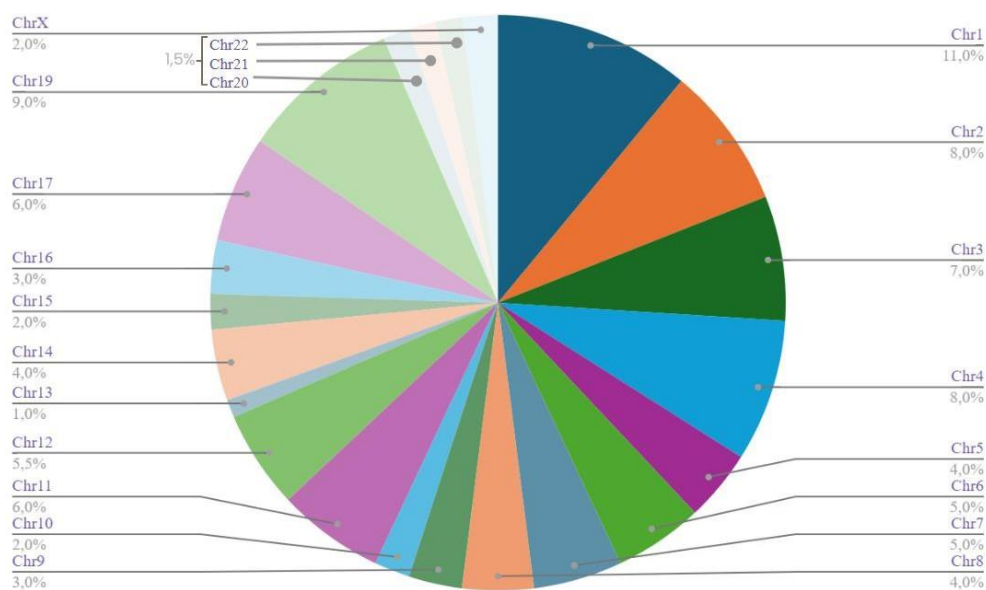
**Figura 4.** Componente Celular de los 200 genes de inflamación

Por otro lado, más de la mitad de genes de estudio (58%) tienen como función molecular la actividad de citoquinas incluyendo *CXCL6; CSF3; CXCL9; CXCL8; CSF1; CX3CL1; CCL7; CCL5; IL12B; CCL2; NDP; TIMP1; CCL17; IL10; CCL24; EDN1; CCL22; IL15; CCL20; CD70; LIF; IL18; OSM; INHBA; CXCL10; IL1A; CXCL11; IL6; IL1B* (Anexo 2.3., Figura 5).



**Figura 5.** Función molecular de los 200 genes de inflamación

Por otro lado, el 11% de genes de inflamación se ubican en el cromosoma 1, seguido del cromosoma 19 (9%) y de los cromosomas 2, 4 y 3 (8%, 8% y 7% respectivamente), según lo obtenido en el Genome Data Viewer (Anexo 3, Figura 6). Asimismo, se observó que ningún gen se ubica en el cromosoma 18, 23 o Y.



**Figura 6.** Distribución de los genes de estudio por cromosoma

### 6.3.Relación entre la ANC y el estado viral

Se observó que, de los 200 genes analizados, el 90% presentaron ANC (179 de 200 genes) en el 70% de pacientes con HCC e infección viral (Anexo 6). En general, la frecuencia de estas alteraciones fue muy baja para los genes relacionados con la inflamación ya que la mayoría de los participantes presentaron entre 1 y 6 amplificaciones o deleciones, observando en la mayoría de casos ausencia de ANC.

Para evaluar si el número de amplificaciones y deleciones difería entre pacientes con y sin infección por HBV/HCV, se utilizó la prueba de Fisher. El análisis reveló que cuatro genes (*CCL24*, *NAMPT*, *SLC7A2* y *SRI*) mostraron frecuencias significativamente diferentes de ANC entre los dos grupos (Tabla 3). La lista completa de genes y los detalles de sus cambios se encuentran en el Anexo 10.

En particular, los genes *CCL24* (quimiocina con motivo CC 24), *NAMPT* (nicotinamida fosforribosiltransferasa) y *SRI* (sorcina) presentaron amplificaciones en individuos sin infección viral ( $p < 0.05$ ), mientras que el gen *SLC7A2* (transportador de aminoácidos

cati3nicos 2) mostr3 deleciones en individuos con infecci3n viral ( $p < 0.05$ ) y ninguna amplificaci3n.

**Tabla 3.** Lista de genes de inflamaci3n con ANC con diferencias significativas en ANCs observadas en pacientes con HCC clasificados seg3n su estado viral (con infecci3n o sin infecci3n viral).

Gen	Nombre del gen	ANC	Virus positivo (n=98)	Virus negativo (n=38)	Total (n=136)	Prueba de Fisher (valor de p)
CCL24	Quimiocina con motivo CC 24	Delecci3n	0	0	0	0,02057
		Amplificaci3n	0	3	3	
		Sin cambios	98	35	133	
NAMPT	Nicotinamida fosforribosiltransferasa	Delecci3n	0	0	0	0,0054
		Amplificaci3n	0	4	4	
		Sin cambios	98	34	132	
SLC7A2	Transportador de amino3cidos cati3nicos 2	Delecci3n	11	0	11	0,03407
		Amplificaci3n	0	0	0	
		Sin cambios	87	38	125	
SRI	Sorcina	Delecci3n	0	0	0	0,0014
		Amplificaci3n	0	5	5	
		Sin cambios	98	33	131	

Adem3s, se analiz3 la frecuencia de amplificaciones y deleciones de los 179 genes con ANC (Anexo 10), en los 136 participantes que presentaron ANC. Los participantes se subdividieron en grupos con y sin infecci3n viral, y se utiliz3 una prueba de Chi cuadrado para evaluar las diferencias entre ambos grupos. Cabe se3alar que casi todos los participantes

presentaron ANC en más de un gen, por lo que el número de frecuencia ya sea de amplificación o delección se incrementó.

La frecuencia de delecciones en los genes asociados con la inflamación fue significativamente mayor en pacientes con HCC y presencia de virus (22%) en comparación con el grupo sin infección viral (5%) ( $p < 0.05$ ). Por otro lado, la frecuencia de amplificaciones en los genes estudiados fue similar en ambos grupos (Tabla 4).

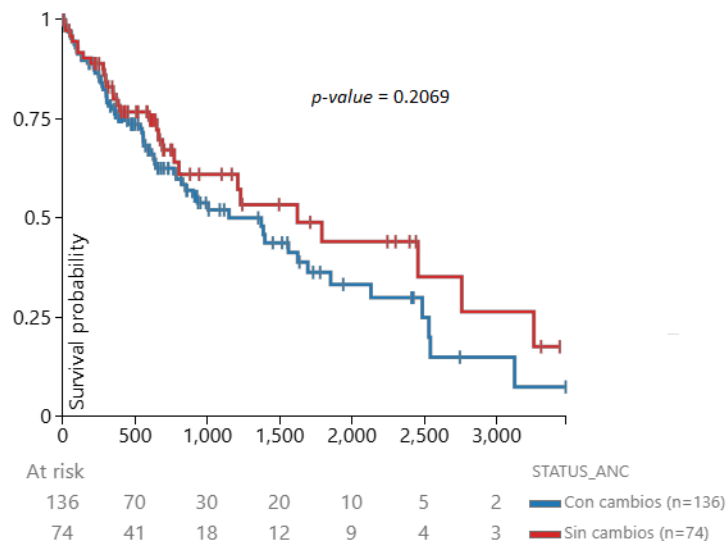
**Tabla 4.** Frecuencia de amplificaciones y delecciones en pacientes con o sin infección viral.

	Virus positivo (n = 599 )	Sin virus (n = 240)	Chi cuadrado (valor de $p$ )	Total ( n = 839)
amp	468 (78.1%)	228 (95%)	0.19	696
del	131 (21.9%)	12 (5%)	0.001069	143

n: número total de eventos de ANC reportados en los 139 participantes, con un máximo de eventos posibles de 17542 (98 x 179) para el grupo con virus positivo y de 6802 (38 x 179) para el grupo sin infección viral.

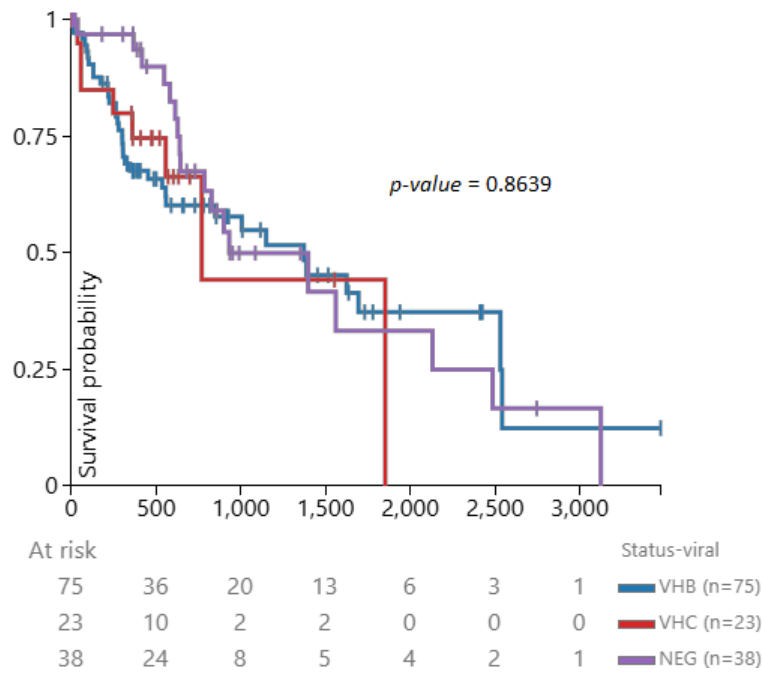
#### **6.4. Asociación entre la ANC y OS en pacientes con infección viral**

La figura 7 presenta el análisis de Kaplan-Meier, generado en Xena Browser a partir de la matriz cargada en esta plataforma (Anexo 8). Esta figura evidenció la probabilidad de supervivencia de los 210 pacientes a lo largo del tiempo, agrupados según la presencia de ANCs (con cambios) en cualquiera de los genes evaluados, o su ausencia (sin cambios). Como era de esperarse, ambas curvas de supervivencia mostraron una tendencia descendente con el tiempo, reflejando una disminución progresiva en la probabilidad de supervivencia en ambos grupos. Aunque el grupo de pacientes sin ANCs mostró una mayor probabilidad de supervivencia en comparación con el grupo con ANCs, esta diferencia no resultó estadísticamente significativa ( $p > 0.05$ ).



**Figura 7.** Plot de Kaplan-Meier representando la probabilidad de sobrevida global respecto al tiempo, en pacientes con HCC agrupados según la presencia de ANC (n=210)

Así mismo, se evaluó la sobrevida, a lo largo del tiempo, de aquellos pacientes con HCC, agrupados por su estado viral, que mostraron ANC en los genes de inflamación (n=179) (Figura 8). A lo largo del tiempo, las curvas de supervivencia de los tres grupos muestran un patrón similar de declive, lo que indicó una disminución en la probabilidad de supervivencia para todos los grupos conforme avanza el tiempo. Es decir, el estado viral (VHB, VHC o negativo) no tuvo un impacto significativo en la supervivencia global de los pacientes ( $p > 0.05$ ).



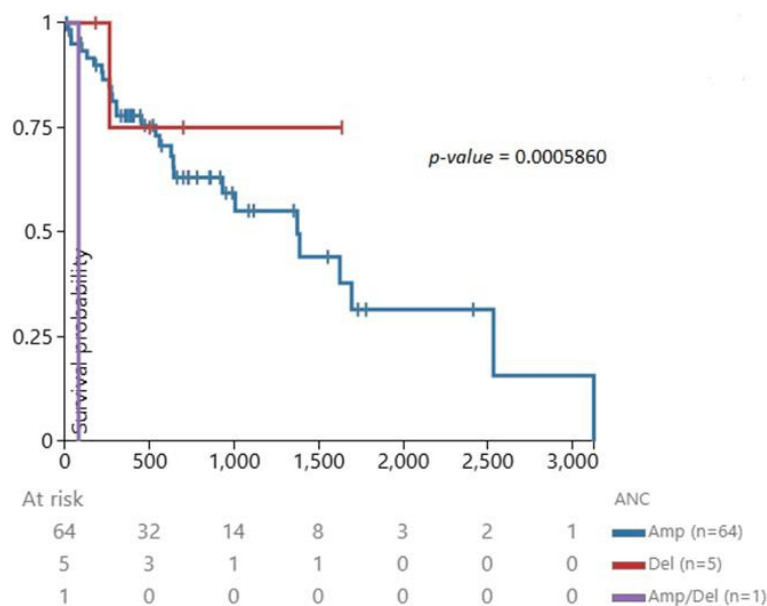
**Figura 8.** Curva de supervivencia de Kaplan-Meier en los 136 pacientes con ANCs, según su estado viral: Hepatitis B, Hepatitis C y sin infección.

Por otra parte, se evaluaron en total 10 plots de KM según el ANC en diferentes rutas enriquecidas de señalización, reguladas por los genes de estudio. Esto se realizó porque, al analizar el ANC por gen individualmente, las bajas frecuencias de ANC sólo habrían incluido a un número reducido de participantes, haciendo que el gráfico de KM no fuera representativo (la metodología se detalla en la sección 6.5.7). Para ello, primero se obtuvo el enriquecimiento funcional basado en el KEGG 2021 el cual mostró 133 rutas con un valor de  $p$  significativo (el Anexo 7), de la cual se consideró las 10 rutas con valores de  $p$  significativos y aquellas que incluían la participación de mayor número de genes. Las rutas fueron corroboradas con la ontología de los genes, específicamente con relación al proceso celular.

Estas 10 rutas fueron las siguientes: Interacción del receptor citocina-citocina, interacción de proteínas virales con citocinas y receptores de citocinas, vía de señalización del TNF, vía de señalización de quimiocinas, vía de señalización JAK-STAT, vía de señalización del receptor

Toll-like, vía de señalización PI3K-Akt, rutas en el cáncer, vía de señalización de NF-kappa B y vía de señalización de IL-17.

Por último, a partir del Anexo 8, se dividieron los participantes y se generaron las matrices correspondientes, presentadas en el Anexo 9, que fueron ingresadas a Xena para plotear la curva de KM. Posteriormente, se descargaron las 10 curvas de Kaplan-Meier (detalladas en el Anexo 12), de las cuales sólo la vía de señalización JAK-STAT (Vía de la Janus Quinasa y el Transductor de Señales y Activador de la Transcripción) tuvo resultados significativos (Figura 9). En ella se observa que los pacientes con deleciones (Del) en los genes reguladores de inflamación tienen una mayor supervivencia global en comparación con aquellos con amplificaciones (Amp).



**Figura 9.** Plot de KM de genes implicados en la vía de señalización JAK-STAT (n=70). *IL10*; *IL15RA*; *CSF3*; *CDKN1A*; *CSF3R*; *IL4R*; *IL15*; *IL10RA*; *LIF*; *OSM*; *OSMR*; *IL6*; *MYC*; *IL2RB*; *IL12B*; *RAF1*; *IL7R*; *IFNAR1* *CCL24*

## VII. Discusión

El HCC suele ser diagnosticado en etapas avanzadas de la enfermedad, donde los pacientes presentan una baja sobrevida [5]. Por ello, es crucial identificar biomarcadores tanto de

diagnóstico como de pronóstico para mejorar la detección temprana y el tratamiento de estos pacientes, respectivamente. En ese sentido, se han descrito algunos marcadores candidatos basados en cambios de expresión de un grupo de genes reguladores de inflamación del HCC [24]. Sin embargo, el estudio de la ANC es menos frecuente, pero podría ser fuente de información relevante para el diagnóstico y tratamiento de pacientes con cáncer.

En vista de ello, este estudio se propuso identificar la presencia de la ANC de genes reguladores de inflamación y su asociación con la OS e infección por VHB/VHC, en pacientes de la cohorte internacional GDC TCGA - LIHC, mediante un análisis secundario de datos genómicos. La asociación de ANC con la OS podría traer nuevos candidatos pronósticos de esta neoplasia, mientras que la relación de ANC con la infección por VHB/VHC, presente hasta en el 70% de casos de HCC [8], podría ayudar a comprender la influencia de estos virus en la progresión tumoral y también en la respuesta al tratamiento. Esto abre la posibilidad de que el estado viral podría ser un factor clave en la personalización de terapias, ayudando a identificar pacientes que podrían beneficiarse de tratamientos específicos según su perfil de infección.

En este contexto, los datos clínicos obtenidos en este estudio indican que la tasa de incidencia de HCC asociado a VHB es mayor en varones que en mujeres, lo cual coincide con las cifras reportadas por la OMS [33]. Según esta organización, del total de 254 millones de personas diagnosticadas con VHB, el 58% son varones. Este hallazgo también ha sido documentado en estudios realizados por Zhang y colaboradores [34].

En la cohorte GDC TCGA-LIHC, se encontró que un 90% genes reguladores de la inflamación presentan ANC. Sin embargo, solo cuatro genes mostraron valores de  $p$  significativos ( $p < 0.05$ ) en la prueba de Fisher, indicando una asociación específica entre

estas alteraciones y la presencia o ausencia del virus. Los genes destacados son: *CCL24* (quimiocina con motivo CC 24), *NAMPT* (nicotinamida fosforribosiltransferasa), *SLC7A2* (transportador de aminoácidos catiónicos 2) y *SRI* (sorcina). Por su parte, el *CCL24* es un factor quimiotáctico que regula la proliferación, migración e invasión de células tumorales, lo que explica su frecuente amplificación. En nuestro estudio, observamos que se amplifica en participantes sin infección viral, lo que sugiere una sobreexpresión. Este hallazgo coincide parcialmente con los resultados de Jin y colaboradores [35], quienes encontraron que en el tejido tumoral el *CCL24* se sobreexpresa y se correlaciona positivamente con la presencia de VHB a diferencia que en el tejido normal. La diferencia radica en que, en nuestro estudio, la amplificación de *CCL24* se asocia positivamente con la ausencia de virus. Esta discrepancia podría explicarse por el hecho de que Jin y colaboradores utilizaron una cohorte distinta a GDC LIHC-TCGA, la cual se utilizó en la presente tesis.

En segundo lugar, encontramos que *NAMPT* está amplificado en individuos sin infección viral, lo que indica que su sobreexpresión puede ser un mecanismo clave para que las células tumorales de HCC mantengan su crecimiento y supervivencia. La sobreexpresión de *NAMPT* asegura que las células cancerosas tengan un suministro constante de NAD<sup>+</sup> (dinucleótido de nicotinamida y adenina) para sostener sus procesos metabólicos, energéticos y de reparación del ADN [36]. Esto convierte a *NAMPT* en un candidato potencial para la terapia dirigida, como ha sido reportado en el estudio previo de Schuster [37], donde sugiere la inhibición de *NAMPT* como un enfoque terapéutico para limitar la capacidad de las células tumorales de mantener su metabolismo acelerado y su resistencia al daño celular.

Asimismo, se evaluó al gen *SRI*, la cual su amplificación revela una sobreexpresión que coincide con hallazgos previos, donde se ha asociado este gen con funciones clave en la migración, invasión, proliferación, resistencia a la apoptosis y una mayor capacidad metastásica [38]. De acuerdo con el estudio de Li y colaboradores [39], la sobreexpresión de *SRI* se ha vinculado con la inhibición de la apoptosis mitocondrial, mediada por la activación de las vías de señalización STAT3 (Transductor de Señales y Activador de la Transcripción 3) y NF- $\kappa$ B (Factor Nuclear kappa B), lo que favorece la proliferación tumoral y la progresión del HCC. Sin embargo, la relación entre estos genes y la infección viral sigue siendo incierta,

dado que tanto *SRI* como *NAMPT*, mencionado anteriormente, se encuentran amplificadas en participantes sin infección viral.

En cuanto a *SLC7A2*, a diferencia de las anteriores, este sufre delección en participantes con presencia del HBV/HCV. Este gen es un transportador de aminoácidos esenciales, principalmente de arginina, lisina y ornitina para ser utilizados en procesos celulares, el metabolismo energético y la síntesis de proteínas. Según el estudio de Xia y colaboradores, la ausencia de *SLC7A2* acelera el desarrollo del carcinoma hepatocelular (HCC) al promover la metástasis y el reclutamiento de células supresoras derivadas de mieloides, las cuales inhiben la respuesta inmune antitumoral [40]. En nuestro estudio, es muy probable que la delección de este gen conduzca a una subexpresión, lo que favorece la progresión del HCC, siendo así un posible biomarcador diagnóstico, ya que sus niveles están directamente relacionados con la agresividad del tumor, ayudando a predecir el curso de la enfermedad. Asimismo, el gen *SLC7A2* se localiza en la región 8p22 del cromosoma, lo cual concuerda con lo reportado anteriormente sobre las pérdidas frecuentes del número de copias génicas en la región 8p [19].

Desde un punto de vista general, es común encontrar amplificaciones en los genes reguladores de inflamación tanto en individuos con infección viral como sin ella, mientras que las delecciones son significativamente mayores en pacientes con infección viral. Estudios como los de Chong et al. [41, 39] y Schulze et al. [42] han reportado amplificaciones frecuentes en genes como *AKT3*, *MYC*, *FGF3* y *TERT*. Sin embargo, estos genes no están exclusivamente relacionados con la regulación de la inflamación, la cual sí se consideró en nuestro estudio. En particular, aunque en nuestra investigación se evaluó el gen *MYC* y se observó su amplificación en la cohorte, la falta de significancia estadística sugiere que este cambio podría tratarse de una alteración común independiente del estado viral.

Por otro lado, la alta frecuencia de delecciones observada en pacientes con infección viral representa un hallazgo novedoso (Tabla 4) y refuerza la evidencia de que el virus induce inestabilidad genómica, un proceso en el que la integración del ADN viral puede alterar genes clave [7,17]. Chong et al. [41] también reportó delecciones en genes como *TP53*, *RBI* y *PTEN*,

siendo *TP53* y *PTEN* ampliamente reconocidos como genes supresores de tumores [43]. En términos generales, un gen supresor de tumores es aquel que regula la división y replicación celular, evitando el crecimiento descontrolado que puede conducir al cáncer. Cuando un gen supresor de tumores muta, su función se pierde o se reduce significativamente, lo que facilita la progresión tumoral [44]. Por lo que en este estudio los genes reguladores de inflamación que presentan deleciones podrían corresponder a genes supresores de tumores, desempeñando un papel relevante en el contexto del HCC. Además, tras realizar el análisis de enriquecimiento funcional, se identificó que uno de los principales caminos en los que están involucrados estos genes es la vía de señalización PI3K-Akt (Anexo 7). Esta vía está estrechamente relacionada con la función de varios genes supresores de tumores, cuya inactivación puede desencadenar la activación descontrolada de esta ruta, contribuyendo significativamente a la progresión tumoral [43].

Con respecto a la sobrevida (OS), en el estudio de Gao y colaboradores [24], se encontró que la sobreexpresión de genes de inflamación como *CCL20*, *IL1B*, *CCL2*, etc (mismos genes evaluados en nuestra investigación), en pacientes con infección viral y HCC, se asocia con una disminución en el tiempo de OS, aunque ellos no evaluaron las ANC. En nuestro análisis, al realizar un plot de KM para evaluar si las ANC influyen en la supervivencia, no se hallaron diferencias significativas entre los grupos con o sin ANC. Posteriormente, al subdividir el grupo de ANC en subgrupos de VHB, VHC y sin infección, el gráfico KM tampoco mostró diferencias significativas. Dado que VHB y VHC son factores etiológicos clave en el desarrollo del HCC, la falta de asociación podría reflejar la heterogeneidad biológica del cáncer hepático, en el que participan múltiples vías y mecanismos moleculares en la carcinogénesis. Además, la OS está influenciada por factores como el tipo de tratamiento recibido y el estadio del tumor, lo que podría explicar la ausencia de una correlación clara en nuestro estudio.

Además, una posible limitación de este estudio es la diferencia en el tipo de genoma de los virus. El VHB es un virus de ADN que se integra en el genoma del huésped, promoviendo inestabilidad genómica y alteraciones en genes clave como *TP53* y *RB1* [45]. Por otro lado, el VHC es un virus de ARN que no se integra en el ADN del huésped, pero induce

carcinogénesis a través de inflamación crónica, estrés oxidativo y activación de vías oncogénicas como NF- $\kappa$ B y JAK-STAT [46]. Estas diferencias pueden influir en la forma en que cada virus afecta la ANC y la progresión tumoral, lo que podría explicar la ausencia de asociaciones significativas en nuestro análisis.

Nuestro estudio mostró que los pacientes con HCC e infección viral que presentan amplificación de genes reguladores de la vía JAK-STAT tienen una menor supervivencia en comparación con aquellos con deleciones en dichos genes, con una media de 655 días y 698 días, respectivamente. Este hallazgo es consistente con la expectativa de que una mayor sobreexpresión de estos genes amplificados reduzca la supervivencia, dado que la vía JAK-STAT está implicada en múltiples procesos oncogénicos en el HCC, como la proliferación celular descontrolada, la evasión de la apoptosis y el escape del sistema inmunológico [47]. Sin embargo, el tamaño reducido del grupo con deleciones (n=5) en comparación con el grupo con amplificaciones (n=68) incrementa la susceptibilidad a resultados aleatorios y limita la representatividad de estos hallazgos en la población general, lo que también afecta su significancia estadística.

Estudios previos con muestras más grandes han demostrado asociaciones más robustas entre infecciones virales y alteraciones genómicas. Por ejemplo, un estudio en *Nature Genetics* incluyó 1,362 individuos y encontró que variantes en IL28B (gen que modula la respuesta inmune antiviral) influyen en la respuesta al tratamiento en VHC [48]. Este ejemplo demuestra que contar con una muestra más amplia permite identificar asociaciones genéticas con mayor precisión, reduciendo la variabilidad estadística y mejorando la representatividad de los hallazgos.

## VIII. Conclusiones

- a. En nuestra investigación se identificó que las ANCs de genes reguladores de inflamación están asociadas con la presencia o ausencia de infecciones virales por VHB y VHC. En particular, genes como *CCL24*, *NAMPT*, *SLC7A2* y *SRI* mostraron patrones específicos de amplificación o delección según el estado viral, sugiriendo que estas alteraciones podrían desempeñar un papel clave en la progresión tumoral y ser consideradas como biomarcadores en el contexto del HCC.
- b. El análisis de la ANC reveló que las delecciones en genes reguladores de la inflamación son más comunes en participantes con infección por VHB/VHC que sin infección, lo que sugiere una posible interacción entre las alteraciones genéticas y la presencia de estos virus en la progresión del carcinoma hepatocelular.
- c. En nuestro estudio aunque no se encontraron diferencias significativas en la OS al comparar pacientes con o sin ANC, se observó que las amplificaciones en genes relacionados con la vía JAK-STAT podrían estar asociadas con una menor OS en pacientes con infección viral. Sin embargo, la interpretación de este hallazgo está limitada por el tamaño reducido del subgrupo, lo que afecta su representatividad estadística y su generalización a la población.

## IX. Limitaciones

Una de las limitaciones en este estudio es el tamaño de la muestra y la cantidad de individuos con información viral disponible, lo que podría haber influido en la interpretación de los resultados. Además, las diferencias en la naturaleza del genoma del VHB y VHC afectan la forma en que estos virus impactan la alteración del número de copias (ANC) y la progresión tumoral, lo que podría explicar la falta de asociaciones significativas observadas. Asimismo, la heterogeneidad biológica del carcinoma hepatocelular (HCC) implica que múltiples mecanismos moleculares pueden estar involucrados en su progresión, lo que limita la capacidad de establecer relaciones directas entre ANC y sobrevida (OS). Adicionalmente, el uso de una única cohorte (GDC TCGA-LIHC) restringe la extrapolación de los hallazgos a otras poblaciones con características genéticas y clínicas distintas. Futuros estudios deberían incluir cohortes internacionales adicionales y profundizar en el análisis de las vías de

señalización asociadas con los genes reguladores de la inflamación para validar y ampliar los hallazgos aquí presentados.

## X. Referencias Bibliográficas

[1] Bray F, Laversanne M, Sung H, Ferlay J, Siegel R L, Soerjomataram I & Jemal A. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36

cancers in 185 countries. *CA: a cancer journal for clinicians*. 2024; 74(3), 229–263. Disponible en: <https://doi.org/10.3322/caac.21834>

[2] Global Cancer Observatory. Top cancer per country, estimated age-standardized mortality rates (World) in 2022 [Internet]. GLOBOCAN; 2022 [citado el 18 de junio de 2024]. Disponible en: <https://gco.iarc.fr/today/en/fact-sheets-populations#countries>

[3] Wang, Z., Qin, H., Liu, S., Sheng, J., & Zhang, X. (2023). Precision diagnosis of hepatocellular carcinoma. *Chinese medical journal*, 136(10), 1155–1165. <https://doi.org/10.1097/CM9.0000000000002641>

[4] Prieto Ortíz JE, Garzón Orjuela N, Sanchez Pardo S, Prieto Ortíz RG, Eslava Schmalbach JH. Hepatocarcinoma: experiencia de la vida real en un centro especializado de Bogotá, Colombia. *Rev Colomb Gastroenterol*. 2022;37(2):163–73. Disponible en: <http://dx.doi.org/10.22516/25007440.823>

[5] Seegobin K, Majeed U, Ritter A, Wylie N, Starr JS, Jones JC, Mody K. Factors affecting time to treatment in hepatocellular cancer. *Journal of Clinical Oncology*. 2020; 38. Disponible en: 10.1200/jco.2020.38.15\_suppl.e19088

[6] Kulik, L., & El-Serag, H. B. Epidemiology and Management of Hepatocellular Carcinoma. *Gastroenterology*, 2019; 156(2), 477–491.e1. Disponible en: <https://doi.org/10.1053/j.gastro.2018.08.065>

- [7] Choi JH, Thung SN. Avances en la clasificación histológica y molecular del carcinoma hepatocelular. *Biomedicines*. 2023; 11(9):2582. <https://doi.org/10.3390/biomedicines11092582>
- [8] Peruhova, M., Banova-Chakarova, S., Miteva, D. G., & Velikova, T. Genetic screening of liver cancer: State of the art. *World journal of hepatology*. 2024; 16(5), 716–730. <https://doi.org/10.4254/wjh.v16.i5.716>
- [9] Centro Nacional de Epidemiología, Prevención y Control de Enfermedades – MINSA. Número de casos de hepatitis B, Perú 2000 – 2023 [Internet]. MINSA; 2020 [citado el 18 de septiembre de 2023]. Disponible en: <https://www.dge.gob.pe/portal/docs/vigilancia/sala/2023/SE03/hepatitisb.pdf>
- [10] Bellido A, Argumanis E, Segura P, Tagle M. Vista de Prevalencia del virus de hepatitis C en donantes de sangre en el Perú 2016 - 2017. *Sociedad de Gastroenterología del Perú*. 2021; 41(3):164-8. Disponible en: <https://revistagastroperu.com/index.php/rgp/article/view/1246/1085>
- [11] Ringelhan M, McKeating JA, Protzer U. Viral hepatitis and liver cancer. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*. 2017; 372(1732): 20160274. Disponible en: <https://doi.org/10.1098/rstb.2016.0274>
- [12] Cheng X, Xia Y, Serti E, Block PD, Chung M, Chayama K, et al. Hepatitis B virus evades innate immunity of hepatocytes but activates macrophages during infection . 2017 [citado el 9 de octubre de 2023];66(6):1779–93. Disponible en: <http://dx.doi.org/10.1002/hep.29348>
- [13] Basit H, Tyagi I, Koirala J. Hepatitis C. [Actualizado el 26 de marzo de 2023]. En: *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing; 2025. Disponible en: <https://www.ncbi.nlm.nih.gov/books/NBK430897/>
- [14] Niu ZS, Niu XJ, Wang WH. Genetic alterations in hepatocellular carcinoma: An update. *World journal of gastroenterology*. 2016; 22(41): 9069–9095. Disponible en: <https://doi.org/10.3748/wjg.v22.i41.9069>

[15] Nevola, R., Tortorella, G., Rosato, V., Rinaldi, L., Imbriani, S., Perillo, P., Mastrocinque, D., La Montagna, M., Russo, A., Di Lorenzo, G., Alfano, M., Rocco, M., Ricozzi, C., Gjeloshi, K., Sasso, F. C., Marfella, R., Marrone, A., Kondili, L. A., Esposito, N., Claar, E. Cozzolino. Gender Differences in the Pathogenesis and Risk Factors of Hepatocellular Carcinoma. *Biology*, 2023, 12(7), 984. <https://doi.org/10.3390/biology12070984>

[16] Sohrab SS, Raj R, Nagar A, Hawthorne S, Paiva-Santos AC, Kamal MA, El-Daly MM, Azhar EI, Sharma A. Chronic Inflammation's Transformation to Cancer: A Nanotherapeutic Paradigm. *Molecules*. 2023; 28(11):4413. <https://doi.org/10.3390/molecules28114413>

[17] Calderaro, J., Couchy, G., Imbeaud, S., Amaddeo, G., Letouzé, E., Blanc, J-F., Laurent, C., Hajji, Y., Azoulay, D., Bioulac-Sage, P., Nault, J-C., Zucman-Rossi, J., Histological Subtypes of Hepatocellular Carcinoma Are Related To Gene Mutations and Molecular Tumour Classification, *Journal of Hepatology*. 2017; 67(4): 727-738. Disponible en: <http://dx.doi.org/10.1016/j.jhep.2017.05.014>

[18] Jiang, Y, Han Q, Zhao H, Zhang J. The Mechanisms of HBV-Induced Hepatocellular Carcinoma. *Journal of hepatocellular carcinoma*. 2021; 8: 435–450. Disponible en: <https://doi.org/10.2147/JHC.S307962>

[19] Ally A, Balasundaram M, Carlsen R, Chuah E, Clarke A., Dha+lla, Ma Y. Comprehensive and Integrative Genomic Characterization of Hepatocellular Carcinoma. *Cell*. 2017; 169(7): 1327–1341. Disponible en doi:10.1016/j.cell.2017.05.046.

[20] Zhao, P., Malik, S., & Xing, S. Epigenetic Mechanisms Involved in HCV-Induced Hepatocellular Carcinoma (HCC). *Frontiers in oncology*, 2021; 11, 677926. Disponible en: <https://doi.org/10.3389/fonc.2021.677926>

[21] Shahrissa A, Tahmasebi-Birgani M, Ansari H, Mohammadi Z, Carloni V, Mohammadi Asl J. The pattern of gene copy number alteration (CNAs) in hepatocellular carcinoma: an in silico analysis. *Molecular cytogenetics*. 2021; 14(1): 33. Disponible en: <https://doi.org/10.1186/s13039-021-00553-2>

[22] Singh N, Baby D, Rajguru J, Patil P, Thakkannavar S, Pujari V. Inflammation and cancer. *Annals of African medicine*. 2019; 18(3): 121–126. Disponible en: [https://doi.org/10.4103/aam.aam\\_56\\_18](https://doi.org/10.4103/aam.aam_56_18)

[23] Refolo MG, Messa C, Guerra V, Carr BI, D'Alessandro R. Inflammatory Mechanisms of HCC Development. *Cancers*. 2020; 12(3): 641. Disponible en: <https://doi.org/10.3390/cancers12030641>

[24] Gao YJ, Li SR, Huang Y. An inflammation-related gene landscape predicts prognosis and response to immunotherapy in virus-associated hepatocellular carcinoma. *Frontiers in oncology*. 2023; 13: 1118152. Disponible en: <https://doi.org/10.3389/fonc.2023.1118152>

[25] El-Nakeep S. Molecular and genetic markers in hepatocellular carcinoma: In silico analysis to clinical validation (current limitations and future promises). *World journal of gastrointestinal pathophysiology*. 2022; 13(1): 1–14. Disponible en: <https://doi.org/10.4291/wjgp.v13.i1.1>

[26] Li Z, Lin Y, Cheng B, Zhang Q, Cai Y. Identification and Analysis of Potential Key Genes Associated With Hepatocellular Carcinoma Based on Integrated Bioinformatics Methods. *Frontiers in Genetics*. 2021;12. Disponible en: <http://dx.doi.org/10.3389/fgene.2021.571231>

[27] Hu X, Zhou J, Zhang Y, Zeng Y, Jie G, Wang S, Yang A, Zhang M. Identifying potential prognosis markers in hepatocellular carcinoma via integrated bioinformatics analysis and biological experiments. *Frontiers in genetics*. 2022; 13: 942454. Disponible en: <https://doi.org/10.3389/fgene.2022.942454>

[28] Bassaganyas L, Pinyol R, Esteban-Fabro R, Torrens L, Torrecilla S, Willoughby CE, Franch-Expósito S, Vila-Casadesús M, Salaverria I, Montal R, Mazzaferro V, Camps J, Sia D, Llovet JM. Copy-Number Alteration Burden Differentially Impacts Immune Profiles and Molecular Features of Hepatocellular Carcinoma. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 2020; 26(23): 6350–6361.

Disponible en: <https://doi.org/10.1158/1078-0432.CCR-20-1497>

[29] Gu, D. L., Chen, Y. H., Shih, J. H., Lin, C. H., Jou, Y. S., & Chen, C. F. Target genes discovery through copy number alteration analysis in human hepatocellular carcinoma. *World journal of gastroenterology*, 2013, 19(47), 8873–8879. <https://doi.org/10.3748/wjg.v19.i47.8873>

[30] Basthi Mohan, P., Lochan, R., & Shetty, S. Biomarker in Hepatocellular Carcinoma. *Indian journal of surgical oncology*. 2024; 15(Suppl 2), 261–268. <https://doi.org/10.1007/s13193-023-01858-x>

[31] Acero Andres. R pubs (2021). Recuperado el 20 de setiembre del 2024: [https://rpubs.com/andresacerol/C3ITAU1#:~:text=La%20funci%C3%B3n%20merge\(\)%20permite,full%20join\)%2C%20entre%20otras.](https://rpubs.com/andresacerol/C3ITAU1#:~:text=La%20funci%C3%B3n%20merge()%20permite,full%20join)%2C%20entre%20otras.)

[32] D'Arrigo, G., Leonardis, D., Abd ElHafeez, S., Fusaro, M., Tripepi, G., & Roumeliotis, S. (2021). Methods to Analyse Time-to-Event Data: The Kaplan-Meier Survival Curve. *Oxidative medicine and cellular longevity*, 2021, 2290120. <https://doi.org/10.1155/2021/2290120>

[33] Organización mundial de la Salud. 2024, Recuperado el 02 de agosto del 2024 de: <https://www.paho.org/es/noticias/10-4-2024-oms-da-alarma-sobre-hepatitis-viricas-que-se-cobran-3500-vidas-al-dia>

[34] Zhang, C. H., Cheng, Y., Zhang, S., Fan, J., & Gao, Q. Changing epidemiology of hepatocellular carcinoma in Asia. *Liver international : official journal of the International Association for the Study of the Liver*, 2022; 42(9), 2029–2041. <https://doi.org/10.1111/liv.15251>

[35] Jin, L., Liu, W. R., Tian, M. X., Jiang, X. F., Wang, H., Zhou, P. Y., Ding, Z. B., Peng, Y. F., Dai, Z., Qiu, S. J., Zhou, J., Fan, J., & Shi, Y. H. CCL24 contributes to HCC malignancy via RhoB- VEGFA-VEGFR2 angiogenesis pathway and indicates poor prognosis. *Oncotarget*, 2017; 8(3), 5135–5148. <https://doi.org/10.18632/oncotarget.14095>

[36] Park JH, Pyun WY, Park HW. Cancer metabolism: Phenotype, signaling and therapeutic targets. *Cells* [Internet]. 2020 [citado el 18 de octubre de 2024];9(10):2308. Disponible en: <https://www.mdpi.com/2073-4409/9/10/2308>

[37] Schuster, S., Penke, M., Gorski, T., Gebhardt, R., Weiss, T. S., Kiess, W., & Garten, A. FK866-induced NAMPT inhibition activates AMPK and downregulates mTOR signaling in hepatocarcinoma cells. *Biochemical and biophysical research communications*, 2015; 458(2), 334–340. <https://doi.org/10.1016/j.bbrc.2015.01.111>

[38] Lu, G., Du, R., Feng, B., Wang, J., Zhang, F., Pei, J., Wang, Y., & Shang, Y. A Novel Gene Signature Associated with Inflammatory Responses and Immune Status Assists in Prognosis and Intervention for Patients with HCC. *Journal of inflammation research*, 2022; 15, 6729–6743. <https://doi.org/10.2147/JIR.S390113>

[39] Li Y, Tian M, Pires Sanches JG, Zhang Q, Hou L, Zhang J. Sorcin inhibits mitochondrial apoptosis by interacting with STAT3 via NF- $\kappa$ B pathway. *Int J Mol Sci* [Internet]. 2024 [citado el 18 de octubre de 2024];25(13):7206. Disponible en: <https://www.mdpi.com/1422-0067/25/13/7206>

[40] Xia, S., Wu, J., Zhou, W., Zhang, M., Zhao, K., Liu, J., Tian, D., & Liao, J. SLC7A2 deficiency promotes hepatocellular carcinoma progression by enhancing recruitment of myeloid-derived suppressors cells. *Cell death & disease*, 2021; 12(6), 570. <https://doi.org/10.1038/s41419-021-03853-y>

[41] Chong, M. L., Knight, J., Peng, G., Ji, W., Chai, H., Lu, Y., Wu, S., Li, P., & Hu, Q.. Integrated exome sequencing and microarray analyses detected genetic defects and underlying pathways of hepatocellular carcinoma. *Cancer genetics*, 2023, 276-277, 30–35. <https://doi.org/10.1016/j.cancergen.2023.06.002>

[42] Schulze, K., Imbeaud, S., Letouzé, E., Alexandrov, L. B., Calderaro, J., Rebouissou, S., Couchy, G., Meiller, C., Shinde, J., Soysouvanh, F., Calatayud, A. L., Pinyol, R., Pelletier, L., Balabaud, C., Laurent, A., Blanc, J. F., Mazzaferro, V., Calvo, F., Villanueva, A., Nault, J. C. Zucman-Rossi, J. Exome sequencing of hepatocellular carcinomas identifies new

mutational signatures and potential therapeutic targets. *Nature genetics*, 2015, 47(5), 505–511. <https://doi.org/10.1038/ng.3252>

[43] Tümen, D., Heumann, P., Gülow, K., Demirci, C. N., Cosma, L. S., Müller, M., & Kandulski, A. Pathogenesis and Current Treatment Strategies of Hepatocellular Carcinoma. *Biomedicines*, 2022, 10(12), 3202. <https://doi.org/10.3390/biomedicines10123202>

[44] Academia Lab. Gen supresor de tumores [Internet]. Academia Lab; [citado 2024 Nov 19]. Disponible en: <https://academia-lab.com/enciclopedia/gen-supresor-de-tumores/>

[45] Yeh, S. H., Li, C. L., Lin, Y. Y., Ho, M. C., Wang, Y. C., Tseng, S. T., & Chen, P. J. Hepatitis B Virus DNA Integration Drives Carcinogenesis and Provides a New Biomarker for HBV-related HCC. *Cellular and molecular gastroenterology and hepatology*, 2023; 15(4), 921–929.

[46] Mitchell, J. K., & McGivern, D. R. Mechanisms of hepatocarcinogenesis in chronic hepatitis C. *Hepatic oncology*, 2014; 1(3), 293–307. <https://doi.org/10.2217/hep.14.7>

[47] Park, H., Lee, S., Lee, J., Moon, H., & Ro, S. W. Exploring the JAK/STAT Signaling Pathway in Hepatocellular Carcinoma: Unraveling Signaling Complexity and

## **Anexos**

Para acceder al Material suplementario ingresar al siguiente enlace:  
[https://github.com/MaryG400/Tesis\\_MaterialSuplementario/tree/main](https://github.com/MaryG400/Tesis_MaterialSuplementario/tree/main)